

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 773 290 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication: 14.05.1997 Bulletin 1997/20 (51) Int. Cl.⁶: **C12N 15/12**, C07K 14/705

(21) Application number: 96116108.0

(22) Date of filing: 08.10.1996

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE

(30) Priority: **09.10.1995 JP 261440**/95 **24.04.1996 JP 102451**/96

(71) Applicant: KOWA COMPANY, LTD. Naka-ku Nagoya-shi Aichi-ken (JP)

(72) Inventors:

- Saito, Yasushi
 Chiba-shi, Chiba (JP)
- Iwasaki, Akio Tsuchiura-shi, Ibaraki (JP)

 Arai, Koichi Urawa-shi, Saitama (JP)

80331 München (DE)

- Yamazaki, Hiroyuki
 Higashimurayama-shi, Tokyo (JP)
- (74) Representative: Wāchtershäuser, Günter, Prof. Dr.
 Patentanwalt,
 Tal 29

Remarks:

The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) Novel LDL receptor analog protein and the gene coding therefor

(57) The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No.5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.

Description

Background of the Invention:

Field of the Invention

The present invention relates to a novel LDL receptor analog protein having a structure similar to that of LDL receptors that are responsible for the homeostasis mechanism of intracellular cholesterol and extensively participates in serum lipid metabolism, which is a critical factor that triggers the onset of arteriosclerosis. The invention also relates to the gene coding for the protein.

2) Description of the Related Art

Abnormality in serum lipid metabolism is one of the most critical risk factors in the onset and progress of arteriosclerosis. Serum lipids, together with apolipoproteins, are transformed into lipoproteins primarily in the liver, secreted therefrom, transported by blood, and taken up by a variety of tissue cells.

Uptake of lipoproteins into cells occurs primarily by the mediation of receptors of respective lipoproteins. It is known that low density lipoproteins (LDL), which are taken into cells by specific membrane receptors, called LDL receptors, are metabolized within the cells and utilized as cell membrane components or similar substances. Detailed analysis of familial hyperchlolesterolemia, which is a genetic disease accompanied by notable hyperchlolesterolemia due to abnormality of LDL receptors, has clarified details of the mechanism of homeostasis achieved by LDL receptors with respect to intracellular cholesterol.

It has been suggested that living bodies have not only LDL receptors but also cell membrane receptors that recognize other lipoproteins. From analyses of WHHL rabbits, which are model animals lacking LDL receptors, it was found that receptors which takes principally apo-E-containing lipoproteins as ligands (remnant receptors) are present in the liver. It is also predicted that there may be HDL receptors whose ligands are high density lipoprotein (HDL). However, to date, details of the structures and functions of these receptors have not yet been elucidated. It has also been known that foaming of macrophages plays an active role in the formation of atherosclerosis, is deeply participated. Macrophages foam by taking up modified LDL—not normal LDL—which have undergone oxidation, acetylation, or glycation. There have recently been discovered receptors to modified LDL which are called scavenger receptors. The scavenger receptors have been identified to be membrane receptors that have a structure completely different from that of LDL receptors.

Recent research using molecular biological techniques has identified the genes of LRP (LDL receptor-associated protein), gp 330, and VLDL receptors. The receptors have been found to have structures very similar to those of LDL receptors. From analyses of these receptors, it is believed that a plurality of lipoprotein receptors are present in living bodies, and that they are closely related to lipid metabolism. LDL receptors studied in detail by Brown and Goldstein [Brown, M.S. and Goldstein, J.L. (1986) Science 232, 34-47] are known to play an important role in the homeostasis of lipoprotein metabolism in vivo, recognizing apo-B-100 and apo-E and taking primarily LDL as their ligands. Also, LRP, which is a macroprotein, has been found to primarily recognize apo-E and to take β -VLDL or chylomicron remnant as a ligand. Moreover, it has been recently reported that LRP takes an α_2 -macroglobulin/protease complex or a plasminogen activator/plasminogen activator inhibitor-1 complex as a ligand, and that LRP is a protein identical to the a2-macroglobulin receptor. When these findings are taken together, LRP is considered to have a wide variety of functions in living bodies [Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H. and Stanley, K.K. (1989) EMBO J. 7(13), 4119-4127; Brown, M.S., Herz, J., Kowal, R.C. and Goldstein, J.L. (1991) Current Opinion in Lipidology 2, 65-72; Herz, J. (1993) Current Opinion in Lipidology 4, 107-113]. The gp 330, which was first identified as an antigen inducing rat Heymann nephritis, has been reported to have a ligand-binding capacity similar to that possessed by CRP ag-macroglobulin receptor [Raychowdhury, R., Niles, J.L., McCluskey, R.T. and Smith, J.A. (1989) Science 244, 1163-1165; Pietromonaco, S., Kerjaschki, D., Binder, S., Ullrich, R. and Farquhar, G. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 1811-1815]. In addition, recently discovered VLDL receptors, which are found to take VLDL as a ligand, are considered to have new functions including fatty acid metabolism, because they are predominantly found in tissues of the heart and muscles though they are rarely found in the liver [Takahashi, S., Kawarabayashi, Y., Nakai, T., Sakai, J. and Yamamoto, T. (1992) Proc. Natl. Acad. Sci. USA 89, 9252-9256].

Functions of these newly found receptors as lipoprotein receptors have been gradually elucidated through detailed in vitro analyses. However, significance of respective receptors in living bodies has mostly been left unknown. In addition, relations to remnant receptors, HDL receptors, etc., which have conventionally been identified or suggested by biochemical techniques, remain unknown. Presently, it is considered that these newly found receptors are products of genes different from those of the latter receptors. Thus, more lipoprotein receptors than originally guessed have become considered to participate in lipoprotein uptake into cells while interacting with each other to thereby function to maintain homeostasis of lipid metabolism in living bodies. However, from structural analyses of the genes of the afore-

mentioned newly-identified receptors, it is predicted that the genes of these receptors that take lipoproteins as ligands are developed from the same gene from which LDL receptors was developed, and thus they are within the same genetic family. This suggests that lipoprotein receptors that have conventionally been proposed may have structures similar to those of LDL receptors.

Accordingly, an object of the present invention is to provide the gene of a novel receptor in the LDL receptor family, as well as a protein coded by the gene.

The present inventors conducted careful studies so as to attain the above object, and found that by using part of rabbit LDL receptor cDNA as a probe there can be obtained a DNA fragment coding for a peptide having a structure similar to that of LDL receptors. Moreover, when using part of the obtained cDNA as a probe, a cDNA fragment having a sequence similar to that of the cDNA can be obtained from the human tissue cDNA library. The present invention was accomplished based on these findings.

Summary of the Invention

5

15

50

The present invention provides DNA having a nucleotide sequence shown by Sequence ID No. 1 or No. 5; an LDL receptor analog protein having an amino acid sequence coded by the DNA; a recombinant vector comprising the DNA and a replicable vector; transformant cells which harbor the recombinant vector; and a method for the production of the LDL receptor analog protein.

Description of Preferred Embodiment

The cDNA of the present invention may be prepared, for example, by the following process.

Briefly, the process includes the following steps. (1) Through the use of rabbit LDL receptor cDNA as a probe, positive clones are screened out of a rabbit liver cDNA library. (2) Recombinant DNA is prepared using the separated positive clones, and a cDNA fragment is cut out of the resultant recombinant DNA through a treatment using a restriction enzyme. The cDNA fragment is integrated into a plasmid vector. (3) Host cells are transformed using the obtained cDNA recombinant vector to thereby obtain transformant cells of the present invention. The obtained transformant cells are incubated so as to obtain a recombinant vector containing a DNA fragment of the present invention. The nucleotide sequence of the DNA fragment of the present invention contained in the resultant recombinant vector is determined. (4) In tissue of a living body, there is detected expression of mRNA indicated by the nucleotide sequence of the cDNA of the present invention by using RNA blot hybridization method. (5) Through use of a rabbit cDNA fragment as a probe, positive clones are screened out of a human tissue cDNA library, and the nucleotide sequence of the clones is determined. (6) A recombinant vector for expression is prepared using the cDNA of the present invention. Through use of the thus-obtained vector, host cells are transformed to thereby obtain the transformants of the present invention. (7) Ligands that are bound to protein expressed by the obtained transformants are detected by ligand blotting.

Each of the above-described steps will next be described.

(1) Screening for positive clones from a rabbit liver cDNA library:

A cDNA library may be prepared by the use of mRNA obtained from rabbit liver, reverse transcriptase, and a suitable vector, e.g., commercially available λgt10 vector.

A cDNA library thus prepared using λgt10 as a vector is subjected to a screening for positive clones by the application of a DNA hybridization method employing a cDNA probe, to thereby separate positive clones [Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) In: Molecular Cloning: A Laboratory Manual, pp 9.47-9.58, Cold Spring Harbor Laboratory Press].

An exemplary cDNA which may be used as a probe is rabbit LDL receptor cDNA. Positive clones may be detected by autoradiography employing a DNA probe labelled with a radioisotope (^{32}P) .

(2) Preparation of a cDNA recombinant vector:

Recombinant vector $\lambda gt10$ phage DNA is extracted from the isolated positive clones and purified. The resultant purified recombinant vector $\lambda gt10$ phage DNA is digested with a restriction enzyme EcoRI, to thereby separate a cDNA fragment from the vector DNA. The obtained cDNA fragment is integrated with a plasmid vector for cloning that has been similarly digested with EcoRI, thereby obtaining a recombinant plasmid vector. An exemplary plasmid vector which may be used is pBluescript II.

(3) Recombinant vector, transformation of host cells using the recombinant vector, and preparation of DNA:

The obtained cDNA recombinant vector is introduced into a variety of host cells that are capable of utilizing the

genetic marker possessed by the recombinant vector, to thereby transform the host cells. Host cells are not particularly limited, with *E. coli* being preferred. For example, a variety of variants of the *E. coli* K12 strain, e.g., HB-101, may be used. In order to introduce the recombinant vector into host cells, a competent cell method may be used in combination with a treatment with calcium.

The thus-obtained transformant cells are cultured in a selective medium in accordance with the genetic marker of the vector. The recombinant vector of the present invention is collected from the cultured cells. The DNA nucleotide sequence of the cDNA contained in the obtained recombinant vector can be determined through use of a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acd. Sci. USA 74, 5463-5467].

(4) RNA blot hybridization:

5

30

40

50

55

The expression in tissue of mRNA, indicated by the nucleotide sequence of the cDNA of the present invention, is detected using RNA blot hybridization.

First, mRNA is prepared using rabbit tissue. Commercially available oligo(dT)cellulose column may be used for the preparation. In order to prepare mRNA from human tissue, there may be used a commercially available nylon membrane on which tissue poly(A)*RNA from a variety of sources is present.

An exemplary probe is the rabbit cDNA obtained in the above-described step (3). mRNA may be detected by autoradiography employing a DNA probe labelled with a radioisotope (32P).

(5) Screening of human tissue cDNA library for positive clones, and determination of nucleotide sequence:

An exemplary human tissue cDNA library which may be used is a commercially available human brain cDNA library.

Screening and nucleotide sequencing of the human brain cDNA library may be performed using a fragment of rabbit cDNA of the present invention as a probe in a manner similar to that used for the aforementioned rabbit liver cDNA library.

(6) Preparation of a recombinant vector for expression and transformation of host cells using the recombinant vector for expression:

In order to prepare an LDL receptor analog protein through use of cDNA of the present invention, the obtained cDNA and a vector for expression are first bonded to each other to thereby create a recombinant vector for expression. Vectors for expression which may be used for bonding are not particularly limited. For example, pBK-CMV may be used.

Host cells are transformed using the thus-obtained recombinant vector for expression, to thereby obtain a transformant cell of the present invention. The obtained transformant cell is cultured so as to obtain cells that are capable of expressing the protein of the invention. Host cells are not particularly limited. For example, CHO cells may be used. In order to introduce the recombinant vector for expression into host cells, a calcium phosphate method may be used.

The thus-prepared transformant cells are incubated in a selective medium in accordance with the genetic marker of the vector, so as to express the LDL receptor analog protein of the present invention.

(7) Ligand analysis of the protein by ligand blotting:

After the resultant transformant cells are incubated, the expressed LDL receptor analog protein is solubilized using a solubilizer, e.g., Triton X-100, to thereby obtain a membrane protein fraction. The fraction is separated using SDS-PAGE, and transferred onto, for example, a nitrocellulose membrane. Using a radio-labelled (125 I) lipoprotein as a probe, the analog protein can be detected by autoradiography. Exemplary lipoproteins which may be used include β -VLDL and LDL.

Examples:

The present invention will next be described in detail by way of example, which should not be construed as limiting the invention.

Example 1:

Preparation of a rabbit liver cDNA library:

From tissue of the liver of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocy-anate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to

thereby obtain purified poly(A)+RNA.

cDNA was synthesized in accordance with a method of Gubler and Hoffman [Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263]. Briefly, cDNA was synthesized employing rabbit liver poly(A)⁺RNA (as a template), a random primer, and moloney murine leukemia virus reverse transcriptase. The synthesized cDNA was transformed into double-stranded DNA using DNA polymerase I, and then subjected to an EcoRI methylase treatment. By the use of T4 DNA polymerase, the DNA was blunt-ended. The blunt-ended DNA was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) using a T4 DNA ligase, and the resultant ligated product was subjected to an additional digestion with EcoRI. cDNA fragments having a size not less than 1 kb were selected by agarose gel electrophoresis, and integrated into the EcoRI-digested site of λgt10 phage DNA using a T4 DNA ligase. The phage DNA was packaged *in vitro*, to thereby establish a rabbit liver cDNA library.

Example 2:

15

Cloning of cDNA of receptors in the rabbit LDL receptor family:

The cDNA library (1,000,000 plaques) prepared in Example 1 was subjected to screening using a plaque hybridization method and employing as a probe a segment of the cDNA obtained from a ligand binding region, the functional region, of the rabbit LDL receptor. Hybridization was performed at 42°C using 5 x SSC, 30% formamide, 1% SDS, 5 x Denhardt's, and 100 μg/ml salmon sperm DNA (ssDNA), followed by washing with 0.3 x SSC/0.1% SDS at 48°C. As a result, several positive clones were obtained. These cDNA clones were separated by performing this plaque hybridization method in a plurality of times. Subsequently, a cDNA fragment of each phage was subcloned into a plasmid vector pBluescript II, and the nucleotide sequence was analyzed using a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acd. Sci. USA 74, 5463-5467]. Based on a putative amino acid sequence, LDL receptors themselves were excluded, and cDNA clones having a sequence very similar to that of LDL receptors were identified. Using these clones as cDNA probes, the cDNA library was screened to thereby obtain overlapping two clones. These were employed as new probes and similar procedure was performed, so as to obtain 5 cDNA clones. The DNA nucleotide sequence determined by these cDNA clones are shown as Sequence ID No. 3. The total length of the sequence was 6961 bp. In the open reading frame of 6639 bp (Sequence ID No. 1) which contained a sequence exhibiting high homology with LDL receptors, there existed on the 5' side an ATG codon which was presumably a translation initiating site and a successive highly hydrophobic sequence consisting of about 30 amino acids. Accordingly, the obtained cDNA was considered to contain the entirety of its length. A putative amino acid sequence is shown as Sequence ID No. 2. The protein consisted of 2213 amino acids. Comparison of the amino acid sequence of the protein with other amino acid sequence data registered at the Genebank, there was a very high similarity to LDL receptors. That is, amino acids 700 - 1,100 in the sequence were very similar to the EGF precursor homology region of LDL receptors, and amino acids 1,100 - 1,640 were also very similar to the ligand binding region of LDL receptors. When the amino acid sequence of the subject protein was compared with other lipoprotein receptor LRP, gp330, and VLDL receptors, similarity was not as high as that observed for LDL receptors. On the C-terminal side of the amino acid sequence of the protein, there was found a highly hydrophobic region which was very similar to the transmembrane region of LDL receptors.

Example 3:

40

From liver tissue and brain tissue of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocyanate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to thereby obtain purified poly(A) $^+$ RNA. The poly(A) $^+$ RNA specimens (10 μ g each) was modified via a glyoxal method, electrophoresed on 1% agarose gel, and transferred onto a nylon membrane.

For human tissue mRNA, commercially available nylon membranes blotted with human tissue poly(A)*RNA from various sources were used.

Using as a probe part of a 32 P-labelled rabbit cDNA of the present invention, hybridization was performed at 42 °C using 50% (rabbit) or 40% (human) formamide, 0.1% SDS, 50 mM phosphate buffer, 5 x Denhardt's, 5 x SSC, and 200 μ g/ml of ssDNA, followed by washing with 0.1 x SSC and 0.1% SDS at 50°C. Autoradiography was performed at $^{-70}$ °C for 2 days in the presence of intensifying screen. As a result, in both rabbit liver tissue and brain tissue, mRNA of about 7 kb was detected as well as mRNA of about 15 kb which was considered to result from alternative splicing or polyadenylation. The size of the mRNA of about 7 kb coincided with that of the rabbit cDNA of the present invention. Also, in human liver tissue and brain tissue, it was confirmed that mRNA having the same size was expressed.

Example 4:

Screening of human brain cDNA library for positive clones and determination of the nucleotide sequence of cDNA fragments

5

10

The human brain cDNA library used in this Example was a commercially obtained cDNA library which was constructed using λ gt10 as a vector. Using partial cDNA of the present invention as a probe, screening of the cDNA library (300,000 plaques) was performed using a plaque hybridization method. Procedures of screening, cloning, and sequencing were as described in Example 2 of the present invention.

As a result of screening of the human brain cDNA library, positive clones containing a DNA fragment of about 3 kb were obtained. Analysis of the nucleotide sequence of part of the cDNA fragment revealed that the fragment was highly homologous to the cDNA of the present invention (Sequence ID No. 4).

Example 5:

15

Cloning of cDNA of receptors in the human LDL receptor family:

A human brain cDNA library was subjected to screening using fragments of the cDNA of the present invention and fragments of the cDNA obtained in Example 4 as probes. Procedures of screening, cloning, and sequencing were as described in Example 2 of the present invention.

Through screening of the human brain cDNA library, two positive clones containing cDNA fragments of about 6 kb and about 3 kb were obtained. When their nucleotide sequence was analyzed, they were identified to be a cDNA clone containing the cDNA nucleotide sequence obtained in Example 4 and a cDNA clone that overlapped therewith. Using part of these cDNAs as probes, procedures similar to those as described above were performed, to thereby obtain another cDNA clone. The DNA nucleotide sequence indicated by these cDNA clones are shown as Sequence ID No. 7. The total length of the sequence was 6,843 bp. There was an open reading frame having a size of 6,642 bp (Sequence ID No. 5). A putative amino acid sequence is shown as Sequence ID No. 6. The protein consisted of 2,214 amino acids. Comparison of the amino acid sequence with that of rabbit protein shown by Sequence ID No. 2 revealed high homology of not less than 94%.

30

35

Example 6:

Creation of cells that express receptors in the rabbit LDL receptor family:

The cDNA as shown by Sequence ID No. 3 was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) by the use of a T4 DNA ligase, and the resultant ligated product was digested with EcoRI. Separately, a vector for expression, pBK-CMV was digested with EcoRI. The aforementioned DNA was ligated to the EcoRI-digested site of the vector using a T4 DNA ligase.

Using the resultant recombinant expression vector in a calcium phosphate method [Chen, C. and H. Okayama (1987) Mol. Cell. Biol. 7, 2745-2752], host cells (CHO-1d1A7) were transformed. The resultant transformants were incubated in a Ham's F-12 selective medium supplemented with 500 µg/ml of G418, and viable cells were separated as LDL receptor analog protein-expressing cells. The cells were incubated further in the aforementioned medium.

Example 7:

45

Ligand analysis of the LDL receptor analog protein by ligand blotting:

The obtained LDL receptor analog protein-expressing cells and control cells were suspended in a buffer solution containing 200 mM Tris-maleic acid (pH 6.5), 2 mM calcium chloride, 0.5 mM PMSF, 2.5 μ M leupeptin, and 1% Triton X-100, to thereby solubilize the membrane protein. Solubilized membrane protein fractions were obtained through centrifugation, and electrophoresed by a 4.5-18% gradient SDS-PAGE. Thereafter, the protein was transferred onto a nitrocellulose membrane.

Incubation was performed in a buffer of 50 mM Tris-HCl (pH 8.0) containing 125 l-labelled β -VLDL (10 μ g/ml), 2 mM calcium chloride, and 5% bovine serum albumin. Autoradiography was performed at room temperature.

A single band of about 250 kDa was detected in membrane protein fractions prepared using the present proteinexpressing cells. This size coincided well with the molecular weight of 248 kDa calculated regarding the amino acid sequence (Sequence ID No. 2) deduced from the cDNA of the present invention. Although a similar band was detected for control cells, the expression level was much lower as compared with the case of the present protein-expressing cells.

Since the protein coded by the cDNA of the present invention is considered to be a novel LDL receptor family recep-

tor, it is expected that through analyses of this protein, details of lipoprotein metabolism mediated by the membrane receptor will be elucidated, and pathology of abnormal lipid metabolism which triggers onset and progress of arteriosclerosis will be clarified.

Sequence ID No. 1

Length of the Sequence: 6639

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Sequence:

5

10

15

20

25

30

40

50

ATGGCGACAC	GGAGCAGCAG	GAGGGAGTCG	CGACTCCCCT	TCCTATTCAC	CCTGGTCGCG	60
CTGCTGCCGC	CCGGGGCTCT	CTGCGAGGTG	TGGACGCGGA	CACTGCACGG	CGGCCGCGCG	120
CCCTTACCCC	AGGAGCGGGG	CTTCCGCGTG	GTGCAGGGCG	ACCCGCGCGA	GCTGCGGCTG	180
TGGGAGCGCG	GGGATGCCAG	GGGGGCGAGC	CGGGCGGACG	AGAAGCCGCT	CCGGAGGAGA	240
CGGAGCGCTG	CCCTGCAGCC	CGAGCCCATC	AAGGTGTACG	GACAGGTCAG	CCTCAATGAT	300
TCCCACAATC	AGATGGTGGT	GCACTGGGCC	GGAGAGAAA	GCAACGTGAT	CGTGGCCTTG	360
GCCCGGGACA	GCCTGGCGTT	GGCCAGGCCC	AGGAGCAGTG	ATGTGTACGT	GTCTTATGAC	420
TATGGAAAAT	CATTCAATAA	GATTTCAGAG	AAATTGAACT	TCGGCGCGGG	AAATAACACA	480
GAGGCTGTGG	TGGCCCAGTT	CTACCACAGC	CCTGCGGACA	ACAAACGGTA	CATCTTCGCA	540
GATGCCTACG	CCCAGTATCT	CTGGATCACG	TTTGACTTCT	GCAACACCAT	CCATGGCTTT	600
TCCATCCCGT	TCCGGGCAGC	TGATCTCCTA	CTCCACAGTA	AGGCCTCCAA	CCTTCTCCTG	660
GGCTTCGACA	GGTCTCACCC	CAACAAGCAG	CTGTGGAAGT	CGGATGATTT	TGGCCAGACC	720
TGGATCATGA	TTCAAGAACA	CGTGAAGTCC	TTTTCTTGGG	GAATTGATCC	CTATGACAAA	780
CCAAACACCA	TCTACATCGA	ACGGCACGAA	CCTTCTGGCT	ACTCCACGGT	TTTCCGAAGT	840
ACAGACTTCT	TCCAGTCCCG	GGAAAACCAG	GAAGTGATCT	TGGAGGAAGT	GAGAGACTTT	900
CAGCTTCGGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTCCACTG	960
CAGTCTTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCGGAAGC	CCATGCGGGC	CGCCCAGTTT	1020
GTTACAAGAC	ATCCTATCAA	CGAATATTAC	ATCGCGGATG	CCTCGGAGGA	CCAGGTGTTT	1080
GTGTGTGTCA	GTCACAGCAA	CAACCGCACC	AACCTCTACA	TCTCGGAGGC	AGAGGGCTTG	1140
AAGTTCTCTC	TGTCCCTGGA	GAACGTGCTC	TACTACACCC	CGGGAGGGGC	CGGCAGTGAC	1200
ACCTTGGTGA	GGTACTTTGC	AAATGAACCG	TTTGCTGACT	TCCATCGTGT	GGAAGGGTTG	1260

	CAGGGAGICI	ACATIGCTAC	TUTGATTAAT	GGTTCTATGA	A PUAUGAGAA	CATGAGATUT	1320
5	GTCATCACCT	TTGACAAAGG	GGGCACCTGG	GAATTTCTGC	AGGCTCCAGC	CTTCACGGGG ·	1380
,	TATGGAGAGA	AAATCAACTG	TGAGCTGTCC	GAGGGCTGTT	CCCTCCACCT	GGCCCAGCGC	1440
	CTCAGCCAGC	TGCTCAACCT	CCAGCTCCGG	AGGATGCCCA	TCCTGTCCAA	GGAGTCGGCG	1500
10	CCTGGCCTCA	TCATTGCCAC	GGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	TCACTACTAT	1620
	ACATGGGGAG	ACCATGGCGG	CATCATCATG	GCCATTGCCC	AAGGCATGGA	AACCAACGAA	1680
15	CTGAAGTACA	GTACCAACGA	AGGGGAGACC	TGGAAAGCCT	TCACCTTCTC	TGAGAAGCCC	1740
	GTGTTTGTGT	ATGGGCTCCT	CACGGAACCC	GGCGAGAAGA	GCACGGTCTT	CACCATCTTT	1800
00	GGCTCCAACA	AGGAGAACGT	GCACAGCTGG	CTCATCCTCC	AGGTCAATGC	CACAGACGCC	1860
20	CTGGGGGTTC	CTTGCACAGA	GAACGACTAC	AAGCTCTGGT	CACCATCTGA	TGAGCGGGGG	1920
	AATGAGTGTT	TGCTTGGACA	CAAGACTGTT	TTCAAACGGA	GGACCCCGCA	CGCCACATGC	1980
25	TTTAACGGAG	AAGACTTTGA	CAGGCCGGTG	GTTGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
·	GACTATGAGT	GTGACTTTGG	CTTCCGGATG	AGTGAAGACT	TGGCATTAGA	GGTGTGTGTT	2100
	CCAGATCCAG	GATTTTCTGG	AAAGTCCTCC	CCTCCAGTGC	CTTGTCCCGT	GGGCTCTACG	2160
30	TACAGGCGAT	CAAGAGGCTA	CCGGAAGATT	TCTGGGGACA	CCTGTAGTGG	AGGAGATGTT	2220
·	GAGGCACGGC	TAGAAGGAGA	GCTGGTCCCC	TGTCCCCTGG	CAGAAGAGAA	CGAGTTCATC	2280
	CTGTACGCCA	CGCGCAAGTC	CATCCACCGC	TATGACCTGG	CTTCCGGAAC	CACGGAGCAG	2340
35	TTGCCCCTCA	CTGGGTTGCG	GGCAGCAGTG	GCCCTGGACT	TTGACTATGA	GCACAACTGC	2400
	CTGTATTGGT	CTGACCTGGC	CTTGGACGTC	ATCCAGCGCC	TCTGTTTGAA	CGGGAGTACA	2460
40	GGACAAGAGG	TGATCATCAA	CTCTGACCTG	GAGACGGTAG	AAGCTTTGGC	TTTTGAACCC	2520
	CTCAGCCAAT	TACTTTACTG	GGTGGACGCA	GGCTTTAAAA	AGATCGAGGT	AGCCAATCCA	2580
	GATGGTGACT	TCCGACTCAC	CGTCGTCAAT	TCCTCGGTGC	TGGATCGGCC	CCGGGCCCTG	2640
45	GTCCTTGTGC	CCCAAGAAGG	GATCATGTTC	TGGACCGACT	GGGGAGACCT	GAAGCCTGGG	2700
	ATTTATCGGA	GCAACATGGA	CGGATCTGCC	GCCTATCGCC	TCGTGTCGGA	GGATGTGAAG	2760
	TGGCCCAATG	GCATTTCCGT	GGACGATCAG	TGGATCTACT	GGACGGATGC	CTACCTGGAC	2820
50	TGCATTGAGC	GCATCACGTT	CAGCGGCCAG	CAGCGCTCCG	TCATCCTGGA	CAGACTCCCG	2880
	CACCCCTATG	CCATTGCTGT	CTTTAAGAAT	GAGATTTACT	GGGATGACTG	GTCACAGCTC	2940

	AGCATATTCC	GAGCTTCTAA	GTACAGCGG	G TCCCAGATG	G AGATTCTGGC	CAGCCAGCTC	3000
5	ACGGGGCTGA	TGGACATGAA	GATCTTCTAC	CAAGGGGAAGA	ACACAGGAAG	CAATGCGTGT	3060
	GTACCCAGGC	CGTGCAGCCT	GCTGTGCCTC	G CCCAGAGCCA	ACAACAGCAA	AAGCTGCAGG	3120
	TGTCCAGATG	GCGTGGCCAG	CAGTGTCCTC	CCTTCCGGGC	ACCTGATGTG	TGACTGCCCT	3180
10	AAGGGCTACG	AGCTGAAGAA	CAACACGTGT	GTCAAAGAAC	AAGACACCTG	TCTGCGCAAC	3240
	CAGTACCGCT	GCAGCAACGG	GAACTGCATC	AACAGCATCT	GGTGGTGCGA	TTTCGACAAC	3300
	GACTGCGGAG	ACÁTGAGCGA	CGAGAAGAAC	TGCCCTACCA	CCATCTGCGA	CCTGGACACC	3360
15	CAGTTCCGTT	GCCAGGAGTC	TGGGACGTGC	ATCCCGCTCT	CCTACAAATG	TGACCTCGAG	3420
	GATGACTGTG	GGGACAACAG	TGACGAAAGG	CACTGTGAAA	TGCACCAGTG	CCGGAGCGAC	3480
20	GAATACAACT	GCAGCTCGGG	CATGTGCATC	CGCTCCTCCT	GGGTGTGCGA	CGGGGACAAC	3540
	GACTGCAGGG	ACTGGTCCGA	CGAGGCCAAC	TGCACAGCCA	TCTATCACAC	CTGTGAGGCC	3600
	TCCAACTTCC	AGTGCCGCAA	CGGGCACTGC	ATCCCCCAGC	GGTGGGCGTG	TGACGGCGAC	3660
25	GCCGACTGCC	AGGATGGCTC	TGATGAGGAT	CCAGCCAACT	GTGAGAAGAA	GTGCAACGGC	3720
	TTCCGCTGCC	CGAACGCCAC	CTGCATTCCC	TCCACCAAGC	ACTGTGACGG	CCTGCACGAT	3780
	TGCTCGGACG	GCTCCGACGA	GCAGCACTGC	GAGCCCCTGT	GTACACGGTT	CATGGACTTC	3840
30	GTGTGTAAGA	ACCGCCAGCA	GTGCCTCTTC	CACTCCATGG	TGTGCGATGG	GATCATCCAG	3900
	TGCCGTGACG	GCTCCGACGA	GGACCCAGCC	TTTGCAGGAT	GCTCCCGAGA	CCCCGAGTTC	3960
o e	CACAAGGTGT	GCGATGAGTT	CGGCTTCCAG	TGTCAGAACG	GCGTGTGCAT	CAGCTTGATC	4020
35	TGGAAGTGCG	ACGGGATGGA	TGACTGCGGG	GACTACTCCG	ACGAGGCCAA	CTGTGAAAAC	4080
	CCCACAGAAG	CCCCCAACTG	CTCCCGCTAC	TTCCAGTTCC	GGTGTGACAA	TGGCCACTGC	4140
10	ATCCCCAACA	GGTGGAAGTG	TGACAGGGAG	AATGACTGTG	GGGACTGGTC	CGACGAGAAG	4200
	GACTGTGGAG	ATTCACATGT	ACTTCCGTCT	ACGACTCCTG	CACCCTCCAC	GTGTCTGCCC	4260
	AATTACTACC	GCTGCGGCGG	GGGGGCCTGC	GTGATAGACA	CGTGGGTTTG	TGACGGGTAC	4320
5	CGAGATTGCG	CAGATGGATC	CGACGAGGAA	GCCTGCCCCT	CGCTCCCCAA	TGTCACTGCC	4380
	ACCTCCTCCC	CCTCCCAGCC	TGGACGATGC	GACCGATTTG	AGTTTGAGTG	CCACCAGCCA	4440
	AAGAAGTGCA	TCCCTAACTG	GAGACGCTGT	GACGGCCATC	AGGATTGCCA	GGATGGCCAG	4500
0	GACGAGGCCA	ACTGCCCCAC	TCACAGCACC	TTGACCTGCA	TGAGCTGGGA	GTTCAAGTGT	4560
	GAGGATGGCG	AGGCCTGCAT	CGTGCTGTCA	GAACGCTGCG	ACGGCTTCCT	GGACTGCTCA	462N

	GATGAGAGCG	ACGAGAAGGC	CTGCAGTGAT	GAGTTAACTG	TATACAAAGT	ACAGAATCTT	4680
5	CAGTGGACAG	CTGACTTCTC	TGGGAATGTC	ACTTTGACCT	GGATGCGGCC	CAAAAAAATG	4740
	CCCTCTGCTG	CTTGTGTATA	CAACGTGTAC	TATAGAGTTG	TTGGAGAGAG	CATATGGAAG	4800
	ACTCTGGAGA	CTCACAGCAA	TAAGACAAAC	ACTGTATTAA	AAGTGTTGAA	ACCAGATACC -	4860
10	ACCTACCAGG	TTAAAGTGCA	GGTTCAGTGC	CTGAGCAAGG	TGCACAACAC	CAATGACTTT	4920
	GTGACCTTGA	GAACTCCAGA	GGGATTGCCA	GACGCCCCTC	AGAACCTCCA	GCTGTCGCTC	4980
	CACGGGGAAG	AGGAAGGTGT	GATTGTGGGC	CACTGGAGCC	CTCCCACCCA	CACCCACGGC	5040
15	CTCATTCGCG	AATACATTGT	AGAGTATAGC	AGGAGTGGTT	CCAAGGTGTG	GACTTCAGAA	5100
	AGGGCTGCTA	GTAACTTTAC	AGAAATAAAG	AACTTGTTGG	TCAACACCCT	GTACACCGTC	5160
	AGAGTGGCTG	CGGTGACGAG	TCGTGGGATA	GGAAACTGGA	GCGATTCCAA	ATCCATTACC	5220
20	ACCGTGAAAG	GAAAAGCGAT	CCCGCCACCA	AATATCCACA	TTGACAACTA	CGATGAAAAT	5280
	TCCCTGAGTT	TTACCCTGAC	CGTGGATGGG	AACATCAAGG	TGAATGGCTA	TGTGGTGAAC	5340
25	CTTTTCTGGG	CATTTGACAC	CCACAAACAA	GAGAAGAAAA	CCATGAACTT	CCAAGGGAGC	5400
	TCAGTGTCCC	ACAAAGTTGG	CAATCTGACA	GCACAGACGG	CCTATGAGAT	TTCCGCCTGG	5460
	GCCAAGACTG	ACTTGGGCGA	TAGTCCTCTG	TCATTTGAGC	ATGTCACGAC	CAGAGGGGTT	5520
30	CGCCCACCTG	CTCCTAGCCT	CAAGGCCAGG	GCTATCAATC	AGACTGCAGT	GGAATGCACC	5580
	TGGACAGGCC	CCAGGAATGT	GGTGTATGGC	ATTTTCTATG	CCACATCCTT	CCTGGACCTC	5640
	TACCGCAACC	CAAGCAGCCT	GACCACGCCG	CTGCACAACG	CAACCGTGCT	CGTCGGTAAG	5700
35	GATGAGCAGT	ATCTGTTTCT	GGTCCGGGTG	GTGATGCCCT	ACCAAGGGCC	GTCCTCGGAC	5760
	TACGTGGTCG	TGAAGATGAT	CCCGGACAGC	AGGCTTCCTC	CCCGCCACCT	GCATGCCGTT	5820
40	CACACCGGCA	AGACCTCGGC	CGTCATCAAG	TGGGAGTCGC	CCTACGACTC	TCCTGACCAG	5880
40	GACCTGTTCT	ATGCGATCGC	AGTTAAAGAT	CTGATACGAA	AGACGGACCG	GAGCTACAAA	5940
	GTCAAGTCCC	GCAACAGCAC	CGTGGAGTAC	ACCCTGAGCA	AGCTGGAGCC	CGGAGGGAAA	6000
45	TACCACGTCA	TTGTGCAGCT	GGGGAACATG	AGCAAAGATG	CCAGTGTGAA	GATCACCACC	6060
	GTTTCGTTAT	CGGCACCCGA	TGCCTTAAAA	ATCATAACAG	AAAATGACCA	CGTCCTTCTC	6120
	TTCTGGAAAA	GTCTAGCTCT	AAAGGAAAAG	TATTTTAACG	AAAGCAGGGG	CTACGAGATA	6180
50	CACATGTTTG	ATAGCGCCAT	GAATATCACC	GCATACCTTG	GGAATACTAC	TGACAATTTC	6240
	TTTAAAATTT	CCAACCTGAA	GATGGGTCAC	AATTACACAT	TCACGGTCCA	GGCACGATGC	6300

	CTTT	TGG	GCA (GCCA	GATC'	TG C	GGGG	AGCC	T GC	CGTG	CTAC	TGT	ATGA′	TGA	GCTG	CGGTCT	6360
5	GGTG	GCG	ATG (CGTC	GGCG	AT G	CAGG	CTGC	C AG	GTCT	ACTG	ATG	TCGC	CGC	CGTG	CTGGTG	6420
•	CCCA	TCC	TGT '	TTCT	GATA	CT G	CTGA	GCCT	G GG	GGTC	GGGT	TTG	CCATO	ССТ	GTAC	ACGAAG	6480
	CATO	CGGA	GGC '	TGCA	GAGC	AG C	TTCA	CCGC	C TT	CGCC	AACA	GCC	ACTAC	CAG	СТСС	AGACTC	6540
10	GGCT	CCG	CCA '	TCTT	CTCC'	тс т	GGGG	ATGA	C TT	GGGG	GAGG	ATG	ATGA	AGA	TGCT	CCTATG	6600
	ATCA	CTG(GAT '	TTTC	GGAC	GA C	GTCC	CCAT	G GT	GATA	GCC						6639
	Sequ	ueno	ce 1	D N	io. :	2											
15	Leng	gth	of	the	Sec	quei	nce:	22	13								
	Туре	e: a	amir	no a	cid												
	Topo	olog	gy:	lin	ear												
20	Mole	ecu.	lar	typ	e:]	Prof	tein										
	Sequ	ueno	ce:														
25	Met	Ala	Thr	Arg	Ser	Ser	Arg	Arg	Glu	Ser	Arg	Leu	Pro	Phe	Leu	Phe	
					5					10					15		
	Thr	Leu	Val	Ala	Leu	Leu	Pro	Pro	Gly	Ala	Leu	Cys	Glu	Val	Trp	Thr	
30				20					25					30			
	Arg	Thr	Leu	His	Gly	Gly	Arg	Ala	Pro	Leu	Pro	Gln	Glu	Arg	Gly	Phe	
			35					40					45				
35	Arg	Val	Val	Gln	Gly	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Trp	Glu	Arg	Gly	
		50					55					60					
40	Asp	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	Lys	Pro	Leu	Arg	Arg	Arg	
,,	65					70					75					80	
	Arg	Ser	Ala	Ala	Leu	G1 n	Pro	Glu	Pro	Ile	Lys	Val	Tyr	Gly	Gln	Val	
<i>15</i>					85					90					95		
	Ser	Leu	Asn	Asp	Ser	His	Asn	Gln	Met	Val	Val	His	Trp	Ala	Gly	Glu	
				100					105					110			
50	Lys :	Ser	Asn	Val	Ιle	Val	Ala	Leu	Ala	Arg	Asp	Ser	Leu	Ala	Leu	Ala	
			115					120					125				

	ALE	110	AI B	261	ser	ASP	vai	ıyr	Val	Ser	TAL	ASP	ıyr	GIY	Lys	Ser
5		130					135					140				
	Phe	Asn	Lys	He	Ser	Glu	Lys	Leu	Asn	Phe	Gly	Ala	Gly	Asn	Asn	Thr
	145					150					155					160
10	Glu	Ala	Val	Val	Ala	Gln	Phe	Туг	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg
					165					170					175	
	Tyr	lle	Phe	Ala	Asp	Ala	Tyr	Ala	Gln	Туг	Leu	Trp	lle	Thr	Phe	Asp
15				180					185					190		
	Phe	Cys	Asn	Thr	Ile	His	Gly	Phe	Ser	Ile	Pro	Phe	Arg	Ala	Ala	Asp
			195					200					205			
20	Leu	Leu	Leu	His	Ser	Lys	Ala	Ser	Asn	Leu	Leu	Leu	Gly	Phe	Asp	Arg
		210					215					220				
25	Ser	His	Pro	Asn	Lys	Gln	Leu	Trp	Lys	Ser	Asp	Asp	Phe	Gly	Gln	Thr
	225					230					235					240
	Trp	lle	Met	Ile	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Trp	Gly	Ιle	Asp
30					245					250					255	
	Pro	Tyr	Asp	Lys	Pro	Asn	Thr	Ile	Туг	Ile	Glu	Arg	His	Glu	Pro	Ser
٠				260					265					270		
35	Gly	Tyr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gln	Ser	Arg	Glu
			275					280					285			
4 0	Asn	Gln	Glu	Val	He	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp
		290					295					300				
•	Lys	Туг	Met	Phe	Ala	Thr	Lys	Val	Val	His	Leu	Leu	Gly	Ser	Pro	Leu
45	305					310					315					320
	Gln	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg
					325					330					335	
50	Ala	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	ΙΙe	Asn	Glu	Туг	Tyr	Ile	Ala
				340					345					350		

	Asp	Ala	Ser	Glu	Asp	Gln	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn
_			355					360					365			
5	Arg	Thr	Asn	Leu	Tyr	lle	Ser	Glu	Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu
		370					375					380				
10	Ser	Leu	Glu	Asn	Val	Leu	Туг	Tyr	Thr	Pro	Gly	Gly	Ala	Gly	Ser	Asp
	385					390					395					400
	Thr	Leu	Val	Arg	Tyr	Phe	Ala	Asn	Glu	Pro	Phe	Ala	Asp	Phe	His	Arg
15					405					410					415	
	Val	Glu	Gly	Leu	Gln	Gly	Val	Туг	Ιlе	Ala	Thr	Leu	Ile	Asn	Gly	Ser
				420					425					430		
20	Met	Asn	Glu	Glu	Asn	Met	Arg	Ser	Val	Ile	Thr	Phe	Asp	Lys	Gly	Gly
			435					440					445			
25	Thr	Trp	Glu	Phe	Leu	Gln	Ala	Pro	Ala	Phe	Thr	Gly	Tyr	Gly	Glu	Lys
25		450					455					460				
	He	Asn	Cys	Glu	Leu	Ser	Glu	Gly	Cys	Ser	Leu	His	Leu	Ala	Gln	Arg
30	465					470					475					480
	Leu	Ser	Gln	Leu	Leu	Asn	Leu	Gln	Leu	Arg	Arg	Met	Pro	Ile	Leu	Ser
					485					490					495	
35	Lys	Glu	Ser	Ala	Pro	Gly	Leu	Ιlе	Ile	Ala	Thr	Gly	Ser	Val	Gly	Lys
				500					505					510		
	Asn	Leu	Ala	Ser	Lys	Thr	Asn	Val	Туг	lle	Ser	Ser	Ser	Ala	Gly	Ala
40			515					520					525			
	Arg	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Туг	Tyr	Thr	Trp	Gly	Asp
45		530					535					540				
	His	Gly	Gly	lle	He	Met	Ala	Ile	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu
	545					550					555					560
50	Leu	Lys	Туг	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	Ala	Phe	Thr	Phe
					565					570					575	

	Ser	Glu	Lys	Pro	Val	Phe	Val	Туг	Gly	Leu	Leu	Thr	Glu	Pro	Gly	Glu
5				580					585					590		
	Lys	Ser	Thr	Val	Phe	Thr	Ιlе	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His
			595					600					605			
10	Ser	Trp	Leu	He	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Pro
		610					615					620				
	Cys	Thr	Glu	Asn	Asp	Туг	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly
15 .	625					630					635					640
	Asn	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro
20					645		•			650					655	
	His	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val
				660					665					670		
25	Ser	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Tyr	Glu	Cys	Asp	Phe	Gly	Phe
			675					680					685			
	Arg	Met	Ser	Glu	Asp	Leu	Ala	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Gly
30		690					695					700				
	Phe	Ser	Gly	Lys	Ser	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Ser	Thr
35	705					710				•	715					720
	Туг	Arg	Arg	Ser	Arg	Gly	Tyr	Arg	Lys	Ile	Ser	Gly	Asp	Thr	Cys	Ser
					725					730					735	
40	Gly	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Cys	Pro
				740					745					750		
	Leu	Ala	Glu	Glu	Asn	Glu	Phe	lle	Leu	Tyr	Ala	Thr	Arg	Lys	Ser	lle
4 5			755					760					765			
	His	Arg	Tyr	Asp	Leu	Ala	Ser	Gly	Thr	Thr	Glu	Gln	Leu	Pro	Leu	Thr
		770					775					780				
50	Gly	Leu	Arg	Ala	Ala	Val	Ala	Leu	Asp	Phe	Asp	Туг	Glu	His	Asn	Cys
	785					790					795					800

	Leu	Tyr	Trp	Ser		Leu	Ala	Leu	Asp		lle	Gln	Arg	Leu		Leu
5					805					810					815	
	Asn	Gly	Ser	Thr	Gly	Gln	Glu	Val	lle	lle	Asn	Ser	Asp	Leu	Glu	Thr
				820					825					830		
10	Val	Glu	Ala	Leu	Ala	Phe	Glu	Pro	Leu	Ser	Gln	Leu	Leu	Туг	Trp	Val
			835					840					845			
	Asp	Ala	Gly	Phe	Lys	Lys	Ιle	Glu	Val	Ala	Asn	Pro	Asp	Gly	Asp	Phe
15		850					855					860				
	Arg	Leu	Thr	Val	Val	Asn	Ser	Ser	Val	Leu	Asp	Arg	Pro	Arg	Ala	Leu
	865					870					875					880
20		Leu	Val	Pro	Gln		Glv	He	Met	Phe		Thr	Asp	Trp	Glv	•
		500	,		885		,			890		****			895	•
	Lou	Lys	Dro	Cly		Tur	Ara	Sar	Acn		۸en	Clv	Ser	Δla		Tvr
25	Leu	Lys			116	ıyı	WI B	261		MCC	пор	u, y	oci	910	A I G	
				900	0.1		v - 1		905	D., a		C1			V - 1	400
00	Arg	Leu			Glu	ASP	Val		Trp	Pro	ASN	uly			Val	ASP
30			915					920					925			
	Asp	Gln	Trp	Пe	Туг	Trp	Thr	Asp	Ala	Туг	Leu	Asp	Cys	lle	Glu	Arg
35		930					935					940	l			
	He	Thr	Phe	Ser	Gly	Gln	Gln	Arg	Ser	Val	He	Leu	Asp	Arg	Leu	Pro
	945	i				950	1				955					960
40	His	Pro	Туг	Ala	Ile	Ala	Val	Phe	Lys	Asn	Glu	He	Туг	Trp	Asp	Asp
					965	j				970	l				975	i
	Trp	Ser	Gln	Leu	Ser	Ile	. Phe	Arg	Ala	Ser	Lys	Туг	Ser	Gly	Ser	Gln
45				980)				985	5				990)	
	Мet	Glu	ı Ile	. Leu	Ala	. Ser	Glr	ı Leu	Thr	Gly	Leu	Met	. Asp	Met	Lys	ille
			995					1000				-	1005			
50	Phe	. Tvr			/ Lvs	s Asr) Thr			· Asr	Ala	. Cvs) Arg	g Pro
		1010			,-		1015		J.,			1020			,	-
		1010	,				1010	,				1020	,			

	Cys	Se	r Le	u Le	и Су	s Le	u Pr	o Ar	g Al	a As	n Asr	s Se	r Ly	s Se	r Cy	s Arg
5	102	5				10	30				103	35				1040
Ţ	Cys	Pro) As	p Gl	y Va	1 A]	a Se	r Se	r Va	l Le	u Pro	Sei	G1;	y As	p Le	u Met
						45				10					10	
10	Cys	Asp	Cy:	s Pr	o Ly	s Gl	у Ту	r Gl	u Le	u Lys	s Asn	Ası	1 Th	r Cy:	s Va	l Lys
				100	60				10	65				10	70	
	Glu	Glu	ı Ası	Thi	r Cy:	s Lei	ı Ar	g Asi	n Gli	n Tyı	Arg	Cys	Sei	r Ası	n Gly	y Asn
15			101					108					108			
	Cys	He	Ası	Sei	- 11	e Trp) Tr	Cys	s Ası	o Phe	e Asp	Asn	Asp	Cys	s Gly	/ Asp
		109					109					110				
20	Met	Ser	Asp	Glu	ı Lys	s Asr	Cys	Pro	Thr	Thr	· Ile	Cys	Asp	. Lei	ı Asp	Thr
	1105					111					111					1120
25	Gln	Phe	Arg	Cys	Glr	ı Glu	Ser	Gly	/ Thr	Cys	Ile	Pro	Leu	ı Ser	Туг	Lys
					112					113					113	
	Cys	Asp	Leu	Glu	Ąsp	Asp	Cys	Gly	Asp	Asn	Ser	Asp	Glu	Arg		
30				114					114					115		·
	Glu	Met	His	Gln	Cys	Arg	Ser	Asp	Glu	Туг	Asn	Cys	Ser			Met
			1155					116					116			
35	Cys	He	Arg	Ser	Ser	Trp	Va]	Cys	Asp	Gly	Asp	Asn	Asp	Cys	Arg	Asp
		1170					1175					1180				-
40	Trp S	Ser	Asp	Glu	Ala	Asn	Cys	Thr	Ala	He	Tyr	His	Thr	Cys	Glu	Ala
40	1185					1190					1195					1200
	Ser #	\sn	Phe	Gln	Cys	Arg	Asn	Gly	His	Cys	He	Pro	Gln	Arg	Тгр	
45					1205					1210					1215	
	Cys A	Sp	Gly	Asp	Ala	Asp	Cys	Gln	Asp	Gly	Ser i	Asp -	Glu	Asp		
				1220					1225					1230		
50	Asn C	ys I	Glu	Lys	Lys	Cys	Asn	Gly	Phe	Arg	Cys 1	Pro i				Cys
			1235					1240					1245			•

	Пe	Pro	Ser	Thr	Lys	His	Cys	Asp	Gly	Leu	His	Asp	Cys	Ser	Asp	Gly
		1250)				1255	5				1260)			
5	Ser	Asp	Glu	Gln	His	Cys	Glu	Pro	Leu	Cys	Thr	Arg	Phe	Met	Asp	Phe
	1265	5				1270)				1275	5				1280
	Val	Cys	Lys	Asn	Arg	Gln	Gln	Cys	Leu	Phe	His	Ser	Met	Val	Cys	Asp
10					1285	i				1290)				1295	i
	Gly	lle	Ile	Gln	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Asp	Pro	Ala	Phe	Ala
15				1300)				1305	5				1310	1	
	Gly	Cys	Ser	Arg	Asp	Pro	Glu	Phe	His	Lys	Val	Cys	Asp	Glu	Phe	Gly
			1315	5				1320)		•		1325	5		
20	Phe	Gln	Cys	Gln	Asn	Gly	Val	Cys	lle	Ser	Leu	lle	Trp	Lys	Cys	Asp
		1330)				133	5				1340)			
	Gly	Met	Asp	Asp	Cys	Gly	Asp	Tyr	Ser	Asp	Glu	Ala	Asn	Cys	Glu	Asn
25	1345	5				135	0				135	5				1360
	Pro	Thr	Glu	Ala	Pro	Asn	Cys	Ser	Arg	Tyr	Phe	Gln	Phe	Arg	Cys	Asp
					1365	5				1370)				1375	5
30	Asn	Gly	His	Cys	Пе	Pro	Asn	Arg	Trp	Ĺys	Cyș	Asp	Arg	Glu	Asn	Asp
•				138	0				138	5				139	0	
35	Cys	Gly	Asp	Trp	Ser	Asp	Glu	Lys	Asp	Cys	Gly	Asp	Ser	His	Val	Leu
33			139	5				1400)				140	5		
	Pro	Ser	Thr	Thr	Pro	Ala	Pro	Ser	Thr	Cys	Leu	Pro	Asn	Tyr	Tyr	Arg
40		141	0				141	5				142	0			
	Cys	Gly	Gly	Gly	Ala	Cys	Val	lle	Asp	Thr	Trp	Val	Cys	Asp	Gly	Tyr
	142	5				143	0				143	5				1440
45	Arg	Asp	Cys	Ala	Asp	Gly	Ser	Asp	Glu	Glu	Ala	Cys	Pro	Ser	Leu	Pro
					144	5				145	0				145	5
	Asn	Val	Thr	Ala	Thr	Ser	Ser	Pro	Ser	Gln	Pro	Gly	Arg	Cys	Asp	Arg
50				146	60				146	55				147	0	

Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe 1525 1530 1535		Phe	Glu	Phe	Glu	Cys	His	Gln	Pro	Lys	Lys	Cys	ΙΙe	Pro	Asn	Trp	Arg
1490 1495 1500 Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys 1505 1510 1510 1515 1520 Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe 1525 1530 1535 Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu 1540 1545 1550 Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly 1555 1560 1565 Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala 1570 1575 1580 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys 1585 1600 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu 1605 1610 1615 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu Gly Val Ile Val Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val	5			147	5				148	0				148	5		
Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys 1505 1510 1515 1520 1505 1510 1515 1520 1525 1530 1535 1535 1520 1535 1520 1535 1520 1535 1540 1545 1550 1550 1540 1545 1550 1565 1550 1565 1560 1565 1550 1565 1570 1575 1580 1570 1575 1580 1585 1590 1585 1600 1585 1600 1615 1635 1630 1635 1630 1635 1630 1635 1630 1635 1630 1635 1630 1635 1630 1635 1630 1645 1645 1645 1645 1645 1655 1660 1645 1655 1650 1655 1660 1635 1650 1655 1650 1645 1645 1645 1655 1650 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1665 1670 1675 1680 1675 1675 1680 1675 1680 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675 1675		Arg	Cys	Asp	Gly	His	Gln	Asp	Cys	Gln	Asp	Gly	Gln	Asp	Glu	Ala	Asn
1505			149	0				149	5				150	0			
Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe 1525 1530 1535 Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu 1540 1540 1545 1550 Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly 1555 1560 1565 25 Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala 1570 1575 1580 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys 30 1585 1590 1595 1600 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu 1605 1610 1615 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val	10	Cys	Pro	Thr	His	Ser	Thr	Leu	Thr	Cys	Met	Ser	Trp	Glu	Phe	Lys	Cys
1525 1530 1535 Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu 1540 1545 1550 Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly 1555 1560 1565 25 Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala 1570 1575 1580 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys 30 1585 1590 1595 1600 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu 1605 1610 1615 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 40 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 45 1650 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680		150	5				151	0				151	5				1520
Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu 1540 1545 1550 Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly 1555 1560 1565 25 Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala 1570 1575 1580 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys 30 1585 1590 1595 1600 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu 1605 1610 1615 25 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 45 1650 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680		Glu	Asp	Gly	Glu	Ala	Cys	He	Val	Leu	Ser	Glu	Arg	Cys	Asp	Gly	Phe
1540 1545 1550 1565 1565 1565 1565 1565 1565 1566 1565 1565 1560 1565 1560 1565 1580 1570 1575 1580 1580 1585 1590 1595 1600 1615 1615 1615 1610 1615 1615 1610 1615 1610 1615 1620 1625 1630 1625 1630 1625 1630 1625 1630 1625 1630 1625 1630 1625 1660 1645 1650 1655 1660 1655 1660 1665 1650 1655 1660 1665 1650 1665 1660	15					152	5				153	0				153	5
Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly 1555 1560 1565 Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala 1570 1575 1580 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys 70 1585 1590 1595 1600 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu 1605 1610 1615 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 45 1650 1655 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val		Leu	Asp	Cys	Ser	Asp	Glu	Ser	Asp	Glu	Lys	Ala	Cys	Ser	Asp	Glu	Leu
Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly 1555					154	0				154	5				155	0	
Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala 1570	20	Thr	Val	Туг	Lys	Val	Gln	Asn	Leu	Gln	Trp	Thr	Ala	Asp	Phe	Ser	Gly
1570				155	5				1560	0				156	5		
Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys 1585 1590 1595 1600 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu 1605 1610 1615 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 1650 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val	25	Asn	Va!	Thr	Leu	Thr	Trp	Met	Arg	Pro	Lys	Lys	Met	Pro	Ser	Ala	Ala
1585 1590 1595 1600 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu 1605 1610 1615 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 45 1650 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val			157	0			٠	157	5				158	0			
The Leu Glu The His See Asn Lys The Asn The Val Leu Lys Val Leu 1605 Lys Pro Asp The The Tyr Gln Val Lys Val Gln Val Gln Cys Leu See 1620 Lys Val His Asn The Asn Asp Phe Val The Leu Arg The Pro Glu Gly 1635 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu See Leu His Gly Glu Glu 45 Glu Gly Val Ile Val Gly His Trp See Pro Pro The His The His Gly 1665 Leu Ile Arg Glu Tyr Ile Val Glu Tyr See Arg See Gly See Lys Val		Cys	Va!	Туг	Asn	Val	Tyr	Tyr	Arg	Val	Val	Gly	Glu	Ser	Пe	Trp	Lys
Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu Lys Val Leu 1605 1610 1615 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1630 1630 1630 1630 1630 1630 1645 1660 1665 1660 1665 1660 1665 1660 1665 1660 1665 1660 1665 1660 1665 1660 1665 1660 1665 1660 1665 1660	30	1585	5				1590)				1595	5				1600
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		Thr	Leu	Glu	Thr	His	Ser	Asn	Lys	Thr	Asn	Thr	Val	Leu	Lys	Val	Leu
Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser 1620 1625 1630 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 45 1650 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val																	
Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly 1635 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 45 1650 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val	35	Lys	Pro	Asp	Thr	Thr	Туг	Gln	Val	Lys	Val	Gln	Val	Gln	Cys	Leu	Ser
1635 1640 1645 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 45 1650 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val					1620)				1625	5				1630)	
Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu 1650 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val	40	Lys	Val	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly
1650 1655 1660 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val				1635	5				1640	1				1645	5		
Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly 1665 1670 1675 1680 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val		Leu	Pro	Asp	Ala	Pro	Gln	Asn	Leu	Gln	Leu	Ser	Leu	His	Gly	Glu	Glu
1665 1670 1675 1680 Leu lie Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val	45		1650)				1655	j				1660)			
Leu lie Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val		Glu	Gly	Val	lle	Val	Gly	His	Тгр	Ser	Pro	Pro	Thr	His	Thr	His	Gly
Leu lie Arg Glu Tyr lie Val Glu Tyr Ser Arg Ser Gly Ser Lys Val		1665	5				1670)				1675	i				1680
1685 1690 1605	50	Leu	Пе	Arg	Glu	Туг	lle	Val	Glu	Туг	Ser	Arg	Ser	Gly	Ser	Lys	Val
1030 1033						1685					1690	}				1695	;

	Trp	Thr	Ser	Glu	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	lle	Lys	Asn	Leu
5				-1700)				1705	5				1710)	
	Leu	Val	Asn	Thr	Leu	Tyr	Thr	Val	Arg	Val	Ala	Ala	Val	Thr	Ser	Arg
			1715	õ				1720)				1725	5		
10	Gly	lle	Gly	Asn	Тгр	Ser	Asp	Ser	Lys	Ser	lle	Thr	Thr	Val	Lys	Gly
		1730)				1735	5				1740)			
	Lys	Ala	He	Pro	Pro	Pro	Asn	lle	His	Ile	Asp	Asn	Tyr	Asp	Glu	Asn
15	1745	5				1750)				1755	5				1760
	Ser	Leu	Ser	Phe	Thr	Leu	Thr	Val	Asp	Gly	Asn	Ile	Lys	Val	Asn	Gly
					176	ō				1770)				1779	5
20	Tyr	Val	Val	Asn	Leu	Phe	Trp	Ala	Phe	Asp	Thr	His	Lys	Gln	Glu	Lys
				178	0				178	5				179	0 -	
25	Lys	Thr	Met	Asn	Phe	Gln	Gly	Ser	Ser	Val	Ser	His	Lys	Val	Gly	Asn
			179	5				180	0				180	5		
	Leu	Thr	Ala	Gln	Thr	Ala	Туг	Glu	Ιle	Ser	Ala	Trp	Ala	Lys	Thr	Asp
30		181	0				181	5				182	0			
	Leu	Gly	Asp	Ser	Pro	Leu	Ser	Phe	Glu	His	Val	Thr	Thr	Arg	Gly	Val
	1825	5				1830)				1835	5				1840
35	Arg	Pro	Pro	Ala	Pro	Ser	Leu	Lys	Ala	Arg	Ala	Ile	Asn	Gln	Thr	Ala
					1845	5				1850)				1855	5
40	Val	Glu	Cys	Thr	Trp	Thr	Gly	Pro	Arg	Asn	Val	Val	Tyr	Gly	He	Phe
				1860)				1869	5				1870)	
	Tyr	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Туг	Arg	Asn	Pro	Ser	Ser	Leu	Thr
45			1875	5				1880)				1889	5		
	Thr	Pro	Leu	His	Asn	Ala	Thr	Val	Leu	Val	Gly	Lys	Asp	Glu	Gln	Tyr
		1890	0				189	5				190	0			
50	Leu	Phe	Leu	Val	Arg	Val	Val	Met	Pro	Tyr	Gln	Gly	Pro	Ser	Ser	Asp
	190	5				191)				191	5				1920

	Туг	Val	Val	Val	Lys	Met	Ile	Pro) Asp	Ser	Arg	Leu	Pro	Pro	Arg	His
5					192	5				193	0				193	5
	Leu	His	Ala	Val	His	Thr	Gly	Lys	Thr	Ser	Ala	Val	Пe	Lys	Trp	Glu
				194	0				194	5				195	D	
10	Ser	Pro	Tyr	Asp	Ser	Pro	Asp	Gln	Asp	Leu	Phe	Туг	Ala	Ile	Ala	Val
			195	5				196	0				196	5		
	Lys	Asp	Leu	Ιle	Arg	Lys	Thr	Asp	Arg	Ser	Tyr	Lys	Val	Lys	Ser	Arg
15		197	0				197	5				198	0			
	Asn	Ser	Thr	Val	Glu	Туг	Thr	Leu	Ser	Lys	Leu	Glu	Pro	Gly	Gly	Lys
20	198	5				199	0				199	5				2000
-0	Tyr	His	Val	Ile	Val	Gln	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ala	Ser	Val
					200	5				201	0				2015	õ
25	Lys	He	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	Ile	Ιle
				202	0				202	5				2030)	
	Thr	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys
30			203	5				204	0				2049	5		
	Glu	Lys	Tyr	Phe	Asn	Glu	Ser	Arg	Gly	Tyr	Glu	lle	His	Met	Phe	Asp
ne		2050)				2055	5				2060)			
75	Ser	Ala	Met	Asn	Пe	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn	Phe
	2065	5				2070)				2075	5				2080
0	Phe	Lys	Ile	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Phe	Thr	Val
					2085	5				2090)				2095	5
	Gln	Ala	Arg	Cys	Leu	Leu	Gly	Ser	Gln	Пе	Cys	Gly	Glu	Pro	Ala	Val
5				2100)				2105	5				2110	İ	
	Leu	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Gly	Gly	Asp	Ala	Ser	Ala	Met	Gln
			2115	j				2120)				2125	i		
o	Ala	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	Пе	Leu	Phe
		2130)				2135	j				2140	1			

	Leu lle Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys
5	2145 2150 2155 2160
Ū	His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr
	2165 2170 2175
10	Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly
	2180 2185 2190
	Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val
15	2195 2200 2205
	Pro Met Val Ile Ala
20	2210
20	Sequence ID No. 3
	Length of the Sequence: 6961
25	Type: nucleic acid
	Strandedness: double
	Topology: linear
30	Molecular type: cDNA to mRNA
	Feature:
25	Name/Key: sig peptide
35	Location: 178261
	Identification method: S
40	Name/Key: mat peptide
	Location: 2626816 Identification method: S
45	Sequence:
	CCGCGAGCCG CACACGTGAC GGCGCCGCGC CGCGCCGCGC
	GGCTGCCCGG AGCCCCGGGA GCGGCGCGCG CGCGGCCCCG GCCCCGCCGC TCGGCCGGC
50	GCGCGCTGCA CATTCTCTCC TGGCGGCGGC GCCACCTGCA GCCGCGTTCG CCCGAACATG 180
	Met

																	i	
_	GCG	ACA	CGG	AGC	AGC	AGG	AGG	GAG	TCG	CGA	СТС	CCC	TTC	СТА	TTC	ACC		228
5	Ala	Thr	Arg	Ser	Ser	Arg	Arg	Glu	Ser	Arg	Leu	Pro	Phe	Leu	Phe	Thr		
				5					10					15				
10	CTG	GTC	GCG	CTG	CTG	CCG	CCC	GGG	GCT	CTC	TGC	GAG	GTG	TGG	ACG	CGG		276
	Leu	Val	Ala	Leu	Leu	Pro	Pro	Gly	Ala	Leu	Cys	Glu	Val	Trp	Thr	Arg		
			20					25					30					
15	ACA .	CTG	CAC	GGC	GGC	CGC	GCG	CCC	TTA	CCC	CAG	GAG	CGG	GGC	TTC	CGC		324
	Thr	Leu	His	Gly	Gly	Arg	Ala	Pro	Leu	Pro	Gln	Glu	Arg	Gly	Phe	Arg		
00		35					4 0					45						
20	GTG	GTG	CAG	GGC	GAC	CCG	CGC	GAG	CTG	CGG	CTG	TGG	GAG	CGC	GGG	GAT		372
	Val	Val	Gln	Gly	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Trp	Glu	Arg	Gly	Asp		
25	50					55			_		60					65		
	GCC	AGG	GGG	GCG	AGC	CGG	GCG	GAC	GAG	AAG	CCG	CTC	CGG	AGG	AGA	CGG		420
	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	Lys	Pro	Leu	Arg	Arg	Arg	Arg		
30					70					7 5					80			
	AGC	GCT	GCC	CTG	CAG	CCC	GAG	CCC	ATC	AAG	GTG	TAC	GGA	CAG	GTC	AGC		468
05	Ser	Ala	Ala	Leu	Gln	Pro	Glu	Pro	lle	Lys	Val	Туг	Gly	Gln	Val	Ser		
35				85					90					95				
	CTC	AAT	GAT	TCC	CAC	AAT	CAG	ATG	GTG	GTG	CAC	TGG	GCC	GGA	GAG	AAA		516
40	Leu	Asn	Asp	Ser	His	Asn	Gln	Met	Val	Val	His	Trp	Ala	Gly	Glu	Lys		
			100					105					110					
						GCC												564
45	Ser			·Ile	Val	Ala			Arg	Asp	Ser			Leu	Ala	Arg		
		115					120					125						
						GTG												612
50		Arg	Ser	Ser	Asp	Val		Val	Ser	Туг		-	Gly	Lys	Ser			
	130					135					140					145		

	AAT	AAG	ATT	TCA	GAG	AAA	TTG	AAC	TTC	GGC	GCG	GGA	AAT	AAC	ACA	GAG	660
5	Asn	Lys	He	Ser	Glu	Lys	Leu	Asn	Phe	Gly	Ala	Gly	Asn	Asn	Thr	Glu	
					150					155					160		
	GCT	GTG	GTG	GCC	CAG	TTC	TAC	CAC	AGC	ССТ	GCG	GAC	AAC	AAA	CGG	TAC	708
10	Ala	Val	Val	Ala	Gln	Phe	Туг	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg	Tyr	
				165					170					175			
	ATC	TTC	GCA	GAT	GCC	TAC	GCC	CAG	TAT	CTC	TGG	ATC	ACG	TTT	GAC	TTC	756
15	lle	Phe	Ala	Asp	Ala	Tyr	Ala	Gln	Tyr	Leu	Trp	He	Thr	Phe	Asp	Phe	
			180					185					190				
	TGC	AAC	ACC	ATC	CAT	GGC	TTT	TCC	ATC	CCG	TTC	CGG	GCA	GCT	GAT	CTC	804
20	Cys	Asn	Thr	lle	His	Gly	Phe	Ser	Ile	Pro	Phe	Arg	Ala	Ala	Asp	Leu	
		195					200					205					
25	CTA	CTC	CAC	AGT	AAG	GCC	TCC	AAC	CTT	СТС	CTG	GGC	TTC	GAC	AGG	TCT	852
	Leu	Leu	His	Ser	Lys	Ala	Ser	Asn	Leu	Leu	Leu	Gly	Phe	Asp	Arg	Ser	
	210					215					2 <u>2</u> 0					225	
30	CAC	CCC	AAC	AAG	CAG	CTG	TGG	AAG	TCG	GAT	GAT	TTT	GGC	CAG	ACC	TGG	900
	His	Pro	Asn	Lys	Gln	Leu	Trp	Lys	Ser	Asp	Asp	Phe	Gly	Gln	Thr	Trp	
					230					235					240		
35	ATC	ATG	ATT	CAA	GAA	CAC	GTG	AAG	TCC	TTT	TCT	TGG	GGA	ATT	GAT	CCC	948
	Ile	Met	He	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Trp	Gly	He	Asp	Pro	
				245					250					255			
10	ТАТ	GAC	AAA	CCA	AAC	ACC	ATC	TAC	ATC	GAA	CGG	CAC	GAA	ССТ	тст	GGC	996
	Tyr	Asp	Lys	Pro	Asn	Thr	He	Туг	lle	Glu	Arg	His	Glu	Pro	Ser	Gly	
4 5	`		260					265					270				
	TAC	TCC	ACG	GTT	TTC	CGA	AGT	ACA	GAC	TTC	TTC	CAG	TCC	CGG	GAA	AAC	1044
	Tyr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gln	Ser	Arg	Glu	Asn	
50		275				_	280		-			285		J			
	CAG		GTG	ATC	TTG	GAG	GAA	GTG	AGA	GAC	TTT		CTT	CGG	GAC	AAG	1092

	Gln	Glu	Val	He	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp	Lys	
5	290					295					300					305	
	TAC	ATG	TTT	GCT	ACA	AAG	GTG	GTG	CAT	CTC	TTG	GGC	AGT	CCA	CTG	CAG	1140
	Tyr	Met	Phe	Ala	Thr	Lys	Val	Val	His	Leu	Leu	Gly	Ser	Pro	Leu	Gln	
10					310					315					320		
	TCT	TCT	GTC	CAG	СТС	TGG	GTC	TCC	TTT	GGC	CGG	AAG	CCC	ATG	CGG	GCC	1188
	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg	Ala	
15				325					330					335			
	GCC	CAG	TTT	GTT	ACA	AGA	CAT	CCT	ATC	AAC	GAA	TAT	TAC	ATC	GCG	GAT	1236
20	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	Ile	Asn	Glu	Tyr	Туг	lle	Ala	Asp	
			340					345					350				
	GCC	TCG	GAG	GAC	CAG	GTG	TTT	GTG	TGT	GTC	AGT	CAC	AGC	AAC	AAC	CGC	1284
25	Ala	Ser	Glu	Asp	Gln	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn	Arg	
		355					360					365					
	ACC	AAC	CTC	TAC	ATC	TCG	GAG	GCA	GAG	GGC	TTG	AAG	TTC	тст	CTG	TCC	1332
30	Thr	Asn	Leu	Туг	He	Ser	Glu	Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu	Ser	
	370					375					380					385	
	CTG	GAG	AAC	GTG	CTC	TAC	TAC	ACC	CCG	GGA	GGG	GCC	GGC	AGT	GAC	ACC	1380
	Leu	Glu	Asn	Val	Leu	Tyr	Tyr	Thr	Pro	Gly	Gly	Ala	Gly	Ser	Asp	Thr	
			-		390					395					400		
40	TTG	GTG	AGG	TAC	TTT	GCA	AAT	GAA	CCG	TTT	GCT	GAC	TTC	CAT	CGT	GTG	1428
	Leu	Val	Arg	Туг	Phe	Ala	Asn	Glu	Pro	Phe	Ala	Asp	Phe	His	Arg	Val	
				405					410					415			
4 5	GAA	GGG	TTG	CAG	GGA	GTC	TAC	ATT	GCT	ACT	CTG	ATT	AAT	GGT	тст	ATG	1476
	Glu	Gly	Leu	Gln	Gly	Val	Tyr	He	Ala	Thr	Leu	Ile	Asn	Gly	Ser	Met	
			420					425					430				
50	AAT	GAG	GAG	AAC	ATG	AGA	тст	GTC	ATC	ACC	TTT	GAC	AAA	GGG	GGC	ACC	1524
	Asn	Glu	Glu	Asn	Met	Arg	Ser	Val	lle	Thr	Phe	Asp	Lys	Gly	Gly	Thr	

		435					440				4	45					
5	TGG	GAA	TTT	CTG	CAG	GCT	CCA	GCC	TTC	ACG	GGG	TAT	GGA	GAG	AAA	ATC	1572
•	Trp	Glu	Phe	Leu	Gln	Ala	Pro	Ala	Phe	Thr	Gly	Туr	Gly	Glu	Lys	lle	
	450					455					460					465	
10	AAC	TGT	GAG	CTG	TCC	GAG	GGC	TGT	TCC	CTC	CAC	CTG	GCC	CAG	CGC	CTC	1620
	Asn	Cys	Glu	Leu	Ser	Glu	Gly	Cys	Ser	Leu	His	Leu	Ala	Gln	Arg	Leu	
					470					475					480		
15	AGC	CAG	CTG	CTC	AAC	СТС	CAG	СТС	CGG	AGG	ATG	CCC	ATC	CTG	TCC	AAG	1668
	Ser	Gln	Leu	Leu	Asn	Leu	Gln	Leu	Arg	Arg	Met	Pro	lle	Leu	Ser	Lys	
				485					490					495			
20	GAG	TCG	GCG	CCT	GGC	CTC	ATC	ATT	GCC	ACG	GGC	TCA	GTG	GGA	AAG	AAC	1716
	Glu	Ser	Ala	Pro	Gly	Leu	Ile	Ile	Ala	Thr	Gly	Ser	Val	Gly	Lys	Asn	
25			500					505					510				
	TTG	GCT	AGC	AAG	ACA	AAC	GTG	TAC	ATC	TCT	AGC	AGT	GCT	GGA	GCC	AGG	1764
	Leu	Ala	Ser	Lys	Thr	Asn	Val	Туг	Ιle	Ser	Ser	Ser	Ala	Gly	Ala	Arg	
30		515					520					525					
	TGG	CGA	GAG	GCA	CTT	CCT	GGA	CCT	CAC	TAC	TAT	ACA	TGG	GGA	GAC	CAT	1812
	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Туг	Туг	Thr	Trp	Gly	Asp	His	
35	530					535					540					545	
	GGC	GGC	ATC	ATC	ATG	GCC	ATT	GCC	CAA	GGC	ATG	GAA	ACC	AAC	GAA	CTG	1860
40	Gly	Gly	lle	Ile	Met	Ala	Ile	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu	Leu	
					550					555	i				560)	
	AAG	TAC	AGT	` ACC	AAC	GAA	GGG	GAG	ACC	TGG	AAA	GCC	TTC	ACC	TT (тст	1908
45	Lys	Туг	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	Ala	Phe	Thr	Phe	Ser	
				565	;				570					575	i		
	GAC	AAC	CCC	GTG	TTT	GTC	ТАТ	GGC	СТС	CTC	ACC	GAA	CCC	GGC	GAC	AAG	1956
50	Glu	Lys	Pro	Val	Phe	. Val	Туг	Gly	Leu	Leu	Thr	Glu	Pro	Gly	Glu	Lys	
			580)				585	j				590)			

	AGC	ACG	GTC	TTC	ACC	ATC	TTT	GGC	TCC	AAC	AAG	GAG	AAC	GTG	CAC	AGC	2004
5	Ser	Thr	Val	Phe	Thr	He	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His	Ser	
		595					600					605					
	TGG	CTC	ATC	СТС	CAG	GTC	AAT	GCC	ACA	GAC	GCC	CTG	GGG	GTT	ССТ	TGC	2052
10	Trp	Leu	Ile	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Рго	Cys	
	610					615					620					625	
	ACA	GAG	AAC	GAC	TAC	AAG	СТС	TGG	TCA	CCA	TCT	GAT	GAG	CGG	GGG	ΑΑΤ	2100
15	Thr	Glu	Asn	Asp	Tyr	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly	Asn	
					630					635					640		
	GAG	TGT	TTG	CTT	GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	AGG	ACC	CCG	CAC	2148
20	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Рго	His	
				645					650					655			
25	GCC	ACA	TGC	TTT	AAC	GGA	GAA	GAC	TTT	GAC	AGG	CCG	GTG	GTT	GTG	TCC	2196
	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val	Ser	
			660					665					670				
30	AAC	TGC	TCC	TGC	ACC	CGG	ĠAG	GAC	TAT	GAG	TGT	GAC	TTT	GGC	TTC	CGG	2244
	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Туг	Glu	Cys	Asp	Phe	Gly	Phe	Arg	
		675					680					685					
35	ATG	AGT	GAA	GAC	TTG	GCA	ATT	GAG	GTG	TGT	GTT	CCA	GAT	CCA	GGA	TTT	2292
	Met	Ser	Glu	Asp	Leu	Ala	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Gly	Phe	
	690					695					700					705	
	TCT	GGA	AAG	TCC	TCC	ССТ	CCA	GTG	CCT	TGT	ccc	GTG	GGC	тст	ACG	TAC	2340
	Ser	Gly	Lys	Ser	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Ser	Thr	Tyr	
15					710					715					720		
	AGG	CGA	TCA	AGA	GGC	TAC	CGG	AAG	ATT	тст	GGG	GAC	ACC	TGT	AGT	GGA	2388
	Arg	Arg	Ser	Arg	Gly	Tyr	Arg	Lys	Ιle	Ser	Gly	Asp	Thr	Cys	Ser	Gly	
50				725					730					735			
	GGA	GAT	GTT	GAG	GCA	CGG	CTA	GAA	GGA	GAG	CTG	GTC	CCC	тст	CCC	CTG	2436

	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Cys	Pro	Leu	
5			740					745					750				
5	GCA	GAA	GAG	AAC	GAG	TTC	ATC	CTG	TAC	GCC	ACG	CGC	AAG	TCC	ATC	CAC	2484
	Ala	Glu	Glu	Asn	Glu	Phe	lle	Leu	Туг	Ala	Thr	Arg	Lys	Ser	Пе	His	
10		755					760	٠				765					
	CGC	TAT	GAC	CTG	GCT	TCC	GGA	ACC	ACG	GAG	CAG	TTG	CCC	CTC	ACT	GGG	2532
	Arg	Tyr	Asp	Leu	Ala	Ser	Gly	Thr	Thr	Glu	Gln	Leu	Pro	Leu	Thr	Gly	
15	770					775					780					785	
	TTC	CGG	GCA	GCA	GTG	GCC	CTG	GAC	TTT	GAC	TAT	GAG	CAC	AAC	TGC	CTG	2580
	Leu	Arg	Ala	Ala	Val	Ala	Leu	Asp	Phe	Asp	Туг	Glu	His	Asn	Cys	Leu	
20					790					795					800		
	TAT	тос	тст	GAC	CTG	GCC	TTG	GAC	GTC	ATC	CAG	CGC	СТС	TGT	TTG	AAC	2628
25	Tyr	Тгр	Ser	Asp	Leu	Ala	Leu	Asp	Val	lle	Gln	Arg	Leu	Cys	Leu	Asn	
				805					810					815			
	GGG	AGT	ACA	GGA	CAA	GAG	GTG	ATC	ATC	AAC	TCT	GAC	CTG	GAG	ACG	GTA	2676
30	Gly	Ser	Thr	Gly	Gln	Glu	Val	Ile	lle	Asn	Ser	Asp	Leu	Glu	Thr	Val	
			820					825				•	830				
	GAA	GCT	TTG	GCT	TTT	GAA	CCC	CTC	AGC	CAA	ATT	CTT	TAC	TGG	GTG	GAC	2724
35	Glu	۸la	Leu	Ala	Phe	Glu	Pro	Leu	Ser	Gln	Leu	Leu	Tyr	Trp	Val	Asp	
		835					840					845					
40	GCA	GGC	TTT	AAA	AAG	ATC	GAG	GTA	GCC	AAT	CCA	GAT	GGT	GAC	TTC	CGA	2772
	Ala	Gly	Phe	Lys	Lys	Ile	Glu	Val	Ala	naA	Pro	Asp	Gly	Asp	Phe	Arg	
	850					855	•				860					865	
45	СТС	ACC	GTC	GTC	TAA	TCC	TCG	GTG	CTG	GAT	CGG	CCC	CGG	GCC	CTG	GTC	2820
	Leu	Thr	Val	Val	Asn	Ser	Ser	Val	Leu	Asp	Arg	Pro	Arg	Ala	Leu	Val	
					870					875					880		
50	CTT	GTG	CCC	CAA	GAA	GGG	ATC	ATG	TTC	TGG	ACC	GAC	TGG	GGA	GAC	CTG	2868
	Leu	Val	Pro	Gln	Glu	Gly	lle	Met	Phe	Trp	Thr	Asp	Trp	Gly	Asp	Leu	

				885					890					895	•		
-	AAG	CCT	GGG	ATT	TAT	CGG	AGC	AAC	ATG	GAC	GGA	тст	GCC	GCC	TAT	CGC	2916
5	Lys	Pro	Gly	lle	Туг	Arg	Ser	Asn	Met	Asp	Gly	Ser	Ala	Ala	Туг	Arg	
			900					905					910				
10	CTC	GTG	TCG	GAG	GAT	GTG	AAG	TGG	ccc	AAT	GGC	TTA	TCC	GTG	GAC	GAT	2964
	Leu	Val	Ser	Glu	Asp	Val	Lys	Trp	Pro	Asn	Gly	Пе	Ser	Val	Asp	Asp	
٠		915					920					925					
15	CAG	TGG	ATC	TAC	TGG	ACG	GAT	GCC	TAC	CTG	GÁC	TGC	АТТ	GAG	CGC	ATC	3012
	Gln	Trp	lle	Туг	Тгр	Thr	Asp	Ala	Tyr	Leu	Asp	Cys	Ile	Glu	Arg	Ile	
	930					935					940					945	
20	ACG	TTC	AGC	GGC	CAG	CAG	CGC	тсс	GTC	ATC	CTG	GAC	AGA	СТС	CCG	CAC	3060
	Thr	Phe	Ser	Gly	Gln	Gln	Arg	Ser	Val	He	Leu	Asp	Arg	Leu	Pro	His	
25					950					955					960		
,	CCC	TAT	GCC	ATT	GCT	GTC	TTT	AAG	AAT	GAG	АТТ	TAC	TGG	GAT	GAC	TGG	3108
	Pro	Tyr	Ala	Ile	Ala	Val	Phe	Lys	Asn	Glu	He	Tyr	Trp	Asp	Asp	Trp	
30				965					970					975			
•	TCA	CAG	СТС	AGC	ATA	TTC	CGA	GCT	TCT	AAG	TAC	AGC	GGG	TCC	CAG	ATG	3156
	Ser	Gln	Leu	Ser	He	Phe	Arg	Ala	Ser	Lys	Туг	Ser	Gly	Ser	Gln	Met	
35			980					985					990				
	GAG	АТТ	CTG	GCC	AGC	CAG	СТС	ACG	GGG	CTG	ATG	GAC	ATG	AAG	ATC	TTC	3204
1 0	Glu	Ile	Leu	Ala	Ser	Gln	Leu	Thr	Gly	Leu	Met	Asp	Met	Lys	He	Phe	
•••		995					100	0				100	5				
	TAC	AAG	GGG	AAG	AAC	ACA	GGA	AGC	AAT	GCG	TGT	GTA	CCC	AGG	CCG	TGC	3252
4 5	Tyr	Lys	Gly	Lys	Asn	Thr	Gly	Ser	Asn	Ala	Cys	Val	Pro	Arg	Pro	Cys	
	101	0				101	5				102	0				1025	
	AGC	CTG	CTG	TGC	CTG	CCC	AGA	GCC	AAC	AAC	AGC	AAA	AGC	TGC	AGG	TGT	3300
50	Ser	Leu	Leu	Cys	Leu	Pro	Arg	Ala	Asn	Asn	Ser	Lys	Ser	Cys	Arg	Cys	
					103	0				103	5				104	0	

	CCA	GAT	GGC	GTG	GCC	AGC	AGT	GTC	CTC	CCT	TCC	GGG	GAC	CTG	ATG	TGT	3348
_	Pro	Asp	Gly	Val	Ala	Ser	Ser	Val	Leu	Pro	Ser	Gly	qzA	Leu	Met	Cys	
5				1045	5				1050)				1055	5		
	GAC	TGC	ССТ	AAG	GGC	TAC	GAG	CTG	AAG	AAC	AAC	ACG	TGT	GTC	AAA	GAA	3396
10	Asp	Cys	Pro	Lys	Gly	Tyr	Glu	Leu	Lys	Asn	Asn	Thr	Cys	Val	Lys	Glu	
			1060)				1065	5				1070)			
	GAA	GAC	ACC	TGT	CTG	CGC	AAC	CAG	TAC	CGC	TGC	AGC	AAC	GGG	AAC	TGC	3444
15	Glu	Asp	Thr	Cys	Leu	Arg	Asn	Gln	Tyr	Arg	Cys	Ser	Asn	Gly	Asn	Cys	
		1075	5				1080)			•	1085	5				
	ATC	AAC	AGC	ATC	TGG	TGG	TGC	GAT	TTC	GAC	AAC	GAC	TGC	GGA	GAC	ATG	3492
20	lle	Asn	Ser	Ile	Trp	Trp	Cys	Asp	Phe	Asp	Asn	Asp	Cys	Gly	Asp	Met	
	1090)				1095	5				1100)				1105	
25	AGC	GAC	GAG	AAG	AAC	TGC	CCT	ACC	ACC	ATC	TGC	GAC	CTG	GAC	ACC	CAG	3540
	Ser	Asp	Glu	Lys	Asn	Cys	Pro	Thr	Thr	lle	Cys	Asp	Leu	Asp	Thr	Gln	
					1110)				1115	5				1120)	
30	TTC	CGT	TGC	CAG	GAG	TCT	GGG	ACG	TGC	ATC	CCG	CTC	TCC	TAC	AAA	TGT	3588
	Phe	Arg	Cys	Gln	Glu	Ser	Gly	Thr	Cys	lle	Pro	Leu	Ser	Tyr	Lys	Cys	
				112	5				1130)				113	5		
35																	
	GAC	CTC	GAG	GAT	GAC	TGT	GGG	GAC	AAC	AGT	GAC	GAA	AGG	CAC	TGT	GAA	3636
															TGT Cys		3636
				Asp					Asn					His			3636
40	Asp	Leu	Glu 114	Asp O	Asp	Cys	Gly	Asp	Asn 5	Ser	Asp	Glu	Arg 115	His O		Glu	3636 3684
40	ASP	Leu CAC	Glu 114 CAG	Asp O TGC	Asp	Cys	Gly GAC	Asp 114 GAA	Asn 5 TAC	Ser	Asp TGC	Glu AGC	Arg 1150 TCG	His O GGC	Cys	Glu TGC	
40 45	ASP	Leu CAC	Glu 114 CAG Gln	Asp O TGC	Asp	Cys	Gly GAC	Asp 114 GAA Glu	Asn 5 TAC	Ser	Asp TGC	Glu AGC	Arg 115 TCG Ser	His O GGC	Cys	Glu TGC	
	Asp ATG Met	CAC His	Glu 114 CAG Gln 5	Asp O TGC Cys	Asp CGG Arg	Cys AGC Ser	Gly GAC Asp	Asp 114 GAA Glu 0	Asn 5 TAC Tyr	Ser AAC Asn	Asp TGC Cys	Glu AGC Ser 116	Arg 115 TCG Ser	His O GGC Gly	Cys ATG Met	Glu TGC	
	ASP ATG Met	CAC His 1150	Glu 1144 CAG Gln 5	Asp 0 TGC Cys	Asp CGG Arg	Cys AGC Ser GTG	Gly GAC Asp 116	Asp 114 GAA Glu O GAC	Asn 5 TAC Tyr GGG	Ser AAC Asn GAC	TGC Cys	AGC Ser 116	Arg 1150 TCG Ser 5	His O GGC Gly AGG	Cys ATG Met	Glu TGC Cys	3684
	ASP ATG Met	CAC His 1155 CGC Arg	Glu 1144 CAG Gln 5	Asp 0 TGC Cys	Asp CGG Arg	Cys AGC Ser GTG	Gly GAC Asp 116 TGC Cys	Asp 114 GAA Glu O GAC	Asn 5 TAC Tyr GGG	Ser AAC Asn GAC	TGC Cys	Glu AGC Ser 116 GAC Asp	Arg 1150 TCG Ser 5	His O GGC Gly AGG	Cys ATG Met	Glu TGC Cys	3684

	Ser	Asp	Glu	Ala	Asn	Cys	Thr	Ala	Ile	Туг	His	Thr	Cys	Glu	Ala	Ser	
5					119	0				119	95				120	0	
	AAC	TTC	CAG	TGC	CGC	AAC	GGG	CAC	TGC	ATC	ccc	CAG	CGG	TGG	GCG	TGT	3828
	Asn	Phe	Gln	Cys	Arg	Asn	Gly	His	Cys	Ιle	Pro	Gln	Arg	Trp	Ala	Cys	
10 -				120	5				121	0				121	5		
	GAC	GGC	GAC	GCC	GAC	TGC	CAG	GAT	GGC	TCT	GAT	GAG	GAT	CCA	GCC	AAC	3876
	Asp	Gly	Asp	Ala	Asp	Cys	Gln	Asp	Gly	Ser	Asp	Glu	Asp	Pro	Ala	Asn	
15			122	0				122	5				123	0			
	TGT	GAG	AAG	AAG	TGC	AAC	GGC	TTC	CGC	TGC	CCG	AAC	GGC	ACC	TGC	ATT	3924
20	Cys	Glu	Lys	Lys	Cys	Asn	Gly	Phe	Arg	Cys	Pro	Asn	Gly	Thr	Cys	Ile	
		123	5				124	0				124	5				
	CCC	TCC	ACC	AAG	CAC	TGT	GAC	GGC	CTG	CAC	GAT	TGC	TCG	GAC	GGC	TCC	3972
25	Pro	Ser	Thr	Lys	His	Cys	Asp	Gly	Leu	His	Asp	Cys	Ser	Asp	Gly	Ser	
	1250)				1255	õ				126	0				1265	
	GAC	GAG	CAG	CAC	TGC	GAG	CCC	CTG	TGT	ACA	CGG	TTC	ATG	GAC	TTC	GTG	4020
30	Asp	Glu	Gln	His	Cys	Glu	Pro	Leu	Cys	Thr	Arg	Phe	Met	Asp	Phe	Val	
	٠				1270)				1275	5				1280)	
35	TGT	AAG	AAC	CGC	CAG	CAG	TGC	СТС	TTC	CAC	тсс	ATG	GTG	TGC	GAT	GGG	4068
33	Cys	Lys	Asn	Arg	Gln	Gln	Cys	Leu	Phe	His	Ser	Met	Val	Cys	Asp	Gly	
				1285	j				1290)				1295	5		
40	ATC	ATC	CAG	TGC	CGT	GAC	GGC	тсс	GAC	GAG	GAC	CCA	GCC	TTT	GCA	GGA	4116
	Пе	He	Gln	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Asp	Pro	Ala	Phe	Ala	Gly	
			1300)				1305	i				1310)			
4 5	TGC	TCC	CGA	GAC	CCC	GAG	TTC	CAC	AAG	GTG	TGC	GAT	GAG	TTC	GGC	TTC	4164
	Cys	Ser	Arg	Asp	Pro	Glu	Phe	His	Lys	Val	Cys	Asp	Glu	Phe	Gly	Phe	
		1315	i				1320)				1325	5				
50	CAG	TGT	CAG	AAC	GGC	GTG	TGC	ATC	AGC	TTG	ATC	TGG	AAG	TGC	GAC	GGG	4212
	Gln	Cys	Gln	Asn	Gly	Val	Cys	Ile	Ser	Leu	Ile	Тгр	Lys	Cys	Asp	Gly	

	1330					1335					1340)			1345					
5	ATG	GAT	GAC	TGC	GGG	GAC	TAC	TCC	GAC	GAG	GCC	AAC	TGT	GAA	AAC	CCC	4260			
5	Met	Asp	Asp	Cys	Gly	Asp	Tyr	Ser	Asp	Glu	Ala	Asn	Cys	Glu	Asn	Pro				
					1350)				1355	5				1360)				
10	ACA	GAA	GCC	CCC	AAC	TGC	TCC	CGC	TAC	TTC	CAG	TTC	CGG	TGT	GAC	AAT	4308			
	Thr	Glu	Ala	Pro	Asn	Cys	Ser	Arg	Туг	Phe	Gln	Phe	Arg	Cys	Asp	Asn				
	1365								1370)				1375	5					
15	GGC	CAC	TGC	ATC	CCC	AAC	AGG	TGG	AAG	TGT	GAC	AGG	GAG	AAT	GAC	TGT	4356			
	Gly	His	Cys	Ile	Pro	Asn	Arg	Trp	Lys	Cys	Asp	Arg	Glu	Asn	Asp	Cys				
	1380							1385	5				1390)						
20	GGG	GAC	TGG	TCC	GAC	GAG	AAG	GAC	TGT	GGA	GÁT	TCA	CAT	GTA	CTT	CCG	4404			
	Gly	Asp	Trp	Ser	Asp	Glu	Lys	Asp	Cys	Gly	Asp	Ser	His	Val	Leu	Pro				
25		1399	5				1400)				140	õ							
	TCT	ACG	ACT	CCT	GCA	CCC	TCC	ACG	TCT	CTG	CCC	AAT	TAC	TAC	CGC	TGC	4452			
	Ser	Thr	Thr	Pro	Ala	Pro	Ser	Thr	Cys	Leu	Pro	Asn	Tyr	Туг	Arg	Cys				
30	1410)				1415	5				1420)				1425				
	GGC	GGG	GGG	GCC	TGC	GTG	ATA	GAC	ACG	TGG	GTT	TGT	GAC	GGG	TAC	CGA	4 500			
	Gly	Gly	Gly	Ala	Cys	Val	Пе	Asp	Thr	Trp	Val	Cys	Asp	Gly	Tyr	Arg				
35					1430)				1435	5 ·				1440)				
	GAT	TGC	GCA	GAT	GGA	TCC	GAC	GAG	GAA	GCC	TGC	CCC	TCG	CTC	CCC	AAT	4548			
40	Asp	Cys	Ala	Asp	Gly	Ser	Asp	Glu	Glu	Ala	Cys	Pro	Ser	Leu	Pro	Asn				
				1445	5				1450)				145	5					
	GTC	ACT	GCC	ACC	TCC	TCC	CCC	TCC	CAG	CCT	GGA	CGA	TGC	GAC	CGA	TTT	4 596			
45	Val	Thr	Ala	Thr	Ser	Ser	Pro	Ser	Gln	Pro	Gly	Arg	Cys	Asp	Arg	Phe				
			146	0				146	5	5			1470							
	GAG	TTT	GAG	TGC	CAC	CAG	CCA	AAG	AAG	TGC	ATC	CCT	AAC	TGG	AGA	CGC	4644			
50	Glu	Phe	Glu	Cys	His	Gln	Pro	Lys	Lys	Cys	Ιlе	Pro	Asn	Trp	Arg	Arg				
	1775						148	0				148	5							

	TGT	GAC	GGC	CAT	CAG	GAT	TGC	CAG	GAT	GGC	CAG	GAC	GAG	GCC	AAC	TGC	4692
5	Cys	Asp	Gly	His	Gln	Asp	Cys	Gln	Asp	Gly	Gln	Asp	Glu	Ala	Asn	Cys	
	1490)				1499	5				1500)				1505	
	CCC	ACT	CAC	AGC	ACC	TTG	ACC	TGC	ATG	AGC	TGG	GAG	TTC	AAG	TGT	GAG	4740
10	Pro	Thr	His	Ser	Thr	Leu	Thr	Cys	Met	Ser	Trp	Glu	Phe	Lys	Cys	Glu	
					1510)				1515	5				1520)	
	GAT	GGC	GAG	GCC	TGC	ATC	GTG	CTG	TCA	GAA	CGC	TGC	GAC	GGC	TTC	CTG	4788
15	Asp	Gly	Glu	Ala	Cys	He	Val	Leu	Ser	Glu	Arg	Cys	Asp	Gly	Phe	Leu	
	1525							1530						1535	5		
	GAC	TGC	TCA	GAT	GAG	AGC	GAC	GAG	AAG	GCC	TGC	AGT	GAT	GAG	ТТА	ACT	4836
20	Asp	Cys	Ser	Asp	Glu	Ser	Asp	Glu	Lys	Ala	Cys	Ser	Asp	Glu	Leu	Thr	
			1540)				1545	5				1550)			
25	GTA	TAC	AAA	GTA	CAG	ААТ	CTT	CAG	TGG	ACA	GCT	GAC	TTC	тст	GGG	AAT	4884
	Val	Туг	Lys	Val	Gln	Asn	Leu	Gln	Тгр	Thr	Ala	Asp	Phe	Ser	Gly	Asn	
		1555					1560	1				1565	₹				
		100	,				1000	,				1000	,				
30	GTC			ACC	TGG	ATG			AAA	AAA	ATG			GCT	GCT	TGT	4932
30		ACT	TTG	ACC Thr			CGG	CCC				CCC	TCT				4932
30		ACT Thr	TTG				CGG Arg	CCC				CCC Pro	TCT				4932
30	Val 1570	ACT Thr)	TTG Leu		Trp	Met 1575	CGG Arg	CCC Pro	Lys	Lys	Met 1580	CCC Pro	TCT Ser	Ala	Ala	Cys 1585	4932 4980
	Val 1570 GTA	ACT Thr) TAC	TTG Leu AAC	Thr	Trp TAC	Met 1575 TAT	CGG Arg 5	CCC Pro	Lys GTT	Lys GGA	Met 1580 GAG	CCC Pro) AGC	TCT Ser	Ala TGG	Ala AAG	Cys 1585 ACT	
35	Val 1570 GTA	ACT Thr) TAC	TTG Leu AAC	Thr GTG	Trp TAC	Met 1575 TAT Tyr	CGG Arg 5	CCC Pro	Lys GTT	Lys GGA	Met 1580 GAG Glu	CCC Pro) AGC	TCT Ser	Ala TGG	Ala AAG	Cys 1585 ACT Thr	
	Val 1570 GTA Val	ACT Thr) TAC Tyr	TTG Leu AAC Asn	Thr GTG	TAC Tyr 1590	Met 1575 TAT Tyr	CGG Arg 5 AGA Arg	CCC Pro GTT Val	Lys GTT Val	Lys GGA Gly 1595	Met 1580 GAG Glu	CCC Pro) AGC Ser	TCT Ser ATA Ile	Ala TGG Trp	Ala AAG Lys 1600	Cys 1585 ACT Thr	
35	Val 1570 GTA Val	Thr) TAC Tyr GAG	TTG Leu AAC Asn	Thr GTG Val	TAC Tyr 1590 AGC	Met 1575 TAT Tyr)	CGG Arg 5 AGA Arg	CCC Pro GTT Val	Lys GTT Val	Lys GGA Gly 1598 ACT	Met 1580 GAG Glu 5	CCC Pro AGC Ser	TCT Ser ATA Ile	Ala TGG Trp GTG	Ala AAG Lys 1600 TTG	Cys 1585 ACT Thr	4980
<i>35</i>	Val 1570 GTA Val	Thr) TAC Tyr GAG	TTG Leu AAC Asn	Thr GTG Val	TAC Tyr 1590 AGC Ser	Met 1575 TAT Tyr)	CGG Arg 5 AGA Arg	CCC Pro GTT Val	Lys GTT Val	Lys GGA Gly 1595 ACT Thr	Met 1580 GAG Glu 5	CCC Pro AGC Ser	TCT Ser ATA Ile	Ala TGG Trp GTG	Ala AAG Lys 1600 TTG Leu	Cys 1585 ACT Thr	4980
35	Val 1570 GTA Val CTG Leu	ACT Thr) TAC Tyr GAG Glu	TTG Leu AAC Asn ACT Thr	Thr GTG Val CAC	TAC Tyr 1590 AGC Ser	Met 1575 TAT Tyr) AAT Asn	CGG Arg 5 AGA Arg AAG Lys	CCC Pro GTT Val ACA Thr	CTT Val AAC Asn 1610	GGA Gly 1595 ACT Thr	Met 1580 GAG Glu 5 GTA Val	CCC Pro) AGC Ser TTA Leu	TCT Ser ATA Ile AAA Lys	TGG Trp GTG Val	Ala AAG Lys 1600 TTG Leu	Cys 1585 ACT Thr AAA Lys	4980
<i>35</i>	Val 1570 GTA Val CTG Leu	ACT Thr) TAC Tyr GAG Glu GAT	TTG Leu AAC Asn ACT Thr	Thr GTG Val CAC His	TAC Tyr 1590 AGC Ser 5	Met 1575 TAT Tyr AAT Asn CAG	CGG Arg AGA Arg AAG Lys	CCC Pro GTT Val ACA Thr	CTT Val AAC Asn 1610 GTG	GGA Gly 1595 ACT Thr CAG	Met 1580 GAG Glu GTA Val	CCC Pro) AGC Ser TTA Leu CAG	TCT Ser ATA Ile AAA Lys	TGG Trp GTG Val 1615 CTG	AAAG Lys 1600 TTG Leu AGC	Cys 1585 ACT Thr AAA Lys	4980 5028
<i>35</i>	Val 1570 GTA Val CTG Leu	ACT Thr) TAC Tyr GAG Glu GAT	TTG Leu AAC Asn ACT Thr	Thr GTG Val CAC His 1605 ACC Thr	TAC Tyr 1590 AGC Ser 5	Met 1575 TAT Tyr AAT Asn CAG	CGG Arg AGA Arg AAG Lys	CCC Pro GTT Val ACA Thr	CTT Val AAC Asn 1610 GTG Val	GGA Gly 1595 ACT Thr CAG	Met 1580 GAG Glu GTA Val	CCC Pro) AGC Ser TTA Leu CAG	TCT Ser ATA Ile AAA Lys	TGG Trp GTG Val 1615 CTG Leu	AAAG Lys 1600 TTG Leu AGC	Cys 1585 ACT Thr AAA Lys	4980 5028
<i>35 40</i>	Val 1570 GTA Val CTG Leu CCA	ACT Thr TAC Tyr GAG Glu GAT Asp	TTG Leu AAC Asn ACT Thr ACC Thr 1620	Thr GTG Val CAC His 1605 ACC Thr	TAC Tyr 1590 AGC Ser TAC Tyr	Met 1575 TAT Tyr AAT Asn CAG	CGG Arg AGA Arg AAG Lys GTT Val	CCC Pro GTT Val ACA Thr AAA Lys 1625	GTT Val AAC Asn 1610 GTG Val	GGA Gly 1595 ACT Thr CAG	Met 1580 GAG Glu GTA Val GTT Val	CCC Pro AGC Ser TTA Leu CAG	TCT Ser ATA Ile AAA Lys TGC Cys 1630	TGG Trp GTG Val 1615 CTG Leu	AAAG Lys 1600 TTG Leu AGC Ser	Cys 1585 ACT Thr AAA Lys AAG Lys	4980 5028

	Val	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly	Leu	
_		1635	5				1640)	•			1645	j				
5	CCA	GAC	GCC	CCT	CAG	AAC	CTC	CAG	CTG	TCG	CTC	CAC	GGG	GAA	GAG	GAA	5172
	Pro	Asp	Ala	Pro	Gln	Asn	Leu	Gln	Leu	Ser	Leu	His	Gly	Glu	Glu	Glu	
10	1656)				1655	5				1660)				1665	
	GGT	GTG	ATT	GTG	GGC	CAC	TGG	AGC	CCT	CCC	ACC	CAC	ACC	CAC	GGC	CTC	5220
	Gly	Val	Ile	Val	Gly	His	Trp	Ser	Pro	Pro	Thr	His	Thr	His	Gly	Leu	
15					1670)				1679	5				1680	0	
	ATT	CGC	GAA	TAC	ATT	GTA	GAG	ТАТ	AGC	AGG	AGT	GGT	TCC	AAG	GTG	TGG	5268
	Ile	Arg	Glu	Tyr	He	Val	Glu	Tyr	Ser	Arg	Ser	Gly	Ser	Lys	Val	Trp	
20				168	5				169	0				169	5		
	ACT	TCA	GAA	AGG	GCT	GCT	AGT	AAC	TTT	ACA	GAA	ATA	AAG	AAC	TTG	TTG	5316
25	Thr	Ser	Glu	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	He	Lys	Asn	Leu	Leu	
25			170	0				170	5				171	0			
	GTC	AAC	ACC	СТG	TAC	ACC	GTC	AGA	GTG	GCT	GCG	GTG	ACG	AGT	· CGT	GGG '	5364
30	Val	Asn	Thr	Leu	Туг	Thr	Val	Arg	Val	Ala	Ala	Val	Thr	Ser	Arg	Gly	
		171	5				172	20				172	5				
	ATA	A GGA	A AAC	TG0	G AGO	GA1	r TC(C AAA	A TC	C AT	г ас	CAC	C GT	G AA	A GG	A AAA	5412
35																y Lys	
	173					173					174					1745	j.
			c ccc	G CC	A CCA	A AA'	TA 7	C CA	CAT	Γ GA	C AA	C TAC	C GA	T GA	A AA	т тсс	5460
40																n Ser	
					179	50				17	55				17	60	
45	СТ	G AG	T TT	r ac	C CT	G AC	C GT	G GA	т GG	G AA	C AT	C AA	G GT	G AA	T GG	C TAT	5508
	Le	u Se	r Phe	e Th	r Lei	u Th	r Va	l As	p G1	y As	n II	e Ly	s Va	l As	n Gl	y Tyr	
				17	65				17	70				17	75		
50	GT	G GT	G AA	с ст	т тт	C TG	G GC	A TT	T GA	C AC	C CA	C AA	A CA	A GA	G AA	G AAA	5556
	۷a	l Va	l As	n Le	u Ph	e Tr	p Al	a Ph	e As	p Th	ır Hi	s Ly	s Gl	n Gl	u Ly	s Lys	

	1780						1785						1790					
5	ACC	ATG	AAC	TTC	CAA	GGG	AGC	TCA	GTG	TCC	CAC	AAA	GTT	GGC	AAT	CTG	5604	
	Thr	Met	Asn	Phe	Gln	Gly	Ser	Ser	Val	Ser	His	Lys	Val	Gly	Asn	Leu		
		179	5				180	0				180	õ					
10	ACA	GCA	CAG	ACG	GCC	TAT	GAG	ATT	TCC	GCC	TGG	GCC	AAG	ACT	GAC	TTG	5652	
	Thr	Ala	Gln	Thr	Ala	Туг	Glu	Ile	Ser	Ala	Trp	Ala	Lys	Thr	Asp	Leu		
	1810					1815	5				1820)		1825				
15	GGC	GAT	AGT	ССТ	CTG	TCA	TTT	GAG	CAT	GTC	ACG	ACC	AGA	GGG	GTT	CGC	5700	
	Gly	Asp	Ser	Pro	Leu	Ser	Phe	Glu	His	Val	Thr	Thr	Arg	Gly	Val	Arg		
										1835	5			~	1840)		
20	CCA	CCT	GCT	ССТ	AGC	СТС	AAG	GCC	AGG	GCT	ATC	AAT	CAG	ACT	GCA	GTG	5748	
	Pro	Pro	Ala	Pro	Ser	Leu	Lys	Ala	Arg	Ala	Πe	Asn	Gln	Thr	Ala	Val		
25	1845						1850							185	5			
	GAA	TGC	ACC	TGG	ACA	GGC	CCC	AGG	AAT	GTG	GŤG	ТАТ	GGC	ATT	TTC	TAT	5796	
	Glu	Cys	Thr	Trp	Thr	Gly	Pro	Arg	Asn	Val	Val	Tyr	Gly	Пе	Phe	Tyr		
30			1860)			1865						1870)		•		
	GCC	ACA	TCC	TTC	CTG	GAC	СТС	TAC	CGC	AAC	CCA	AGC	AGC	СТG	ACC	ACG	5844	
	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Туг	Arg	Asn	Pro	Ser	Ser	Leu	Thr	Thr		
35		1875	5				1880)				1885	5					
	CCG	CTG	CAC	AAC	GCA	ACC	GTG	СТС	GTC	GGT	AAG	GAT	GAG	CAG	TAT	CTG	5892	
	Pro	Leu	His	Asn	Ala	Thr	Val	Leu	Val	Gly	Lys	Asp	Glu	Gln	Tyr	Leu		
10	1890)				1895	5				1900)				1905		
	TTT	CTG	GTC	CGG	GTG	GTG	ATG	CCC	TAC	CAA	GGG	CCG	TCC	TCG	GAC	TAC	5940	
15	Phe	Leu	Val	Arg	Val	Val	Met	Pro	Туг	Gln	Gly	Pro	Ser	Ser	Asp	Tyr		
					1910)				1915	5				1920			
	GTG	GTC	GTG	AAG	ATG	ATC	CCG	GAC	AGC	AGG	СТТ	ССТ	CCC	CGG	CAC	CTG	5988	
ro .	Val	Val	Val	Lys	Met	Ιle	Pro	Asp	Ser	Arg	Leu	Pro	Pro	Arg	His	Leu		
	Val Val Val Lys Met 1925								1930)				1935				

	CAT	GCC	GTT	CAC	ACC	GGC	AAG	ACC	TCG	GCC	GTC	ATC	AAG	TGG	GAG	TCG	6036	
5	His	Ala	Val	His	Thr	Gly	Lys	Thr	Ser	Ala	Val	Пe	Lys	Trp	Glu	Ser		
•			194	0				194	5				1950)				
	CCC	TAC	GAC	TCT	CCT	GAC	CAG	GAC	CTG	TTC	TAT	GCG	ATC	GCA	GTT	AAA	6084	
10	Pro	Tyr	Asp	Ser	Pro	Asp	G1 n	Asp	Leu	Phe	Tyr	Ala	Ile	Ala	Val	Lys		
		195	5				1960					196	5					
	GAT	CTG	ATA	CGA	AAG	ACG	GAC	CGG	AGC	TAC	AAA	GTC	AAG	TCC	CGC	AAC	6132	
15	Asp	Leu	He	Arg	Lys	Thr	Asp	Arg	Ser	Tyr	Lys	Val	Lys	Ser	Arg	Asn		
	1970)				1975	j			1980						1985		
	AGC	ACC	GTG	GAG	TAC	ACC	CTG	AGC	AAG	CTG	GAG	CCC	GGA	GGG	AAA	TAC	6180	
20	Ser	Thr	Val	Glu	Tyr	Thr	Leu	Ser	Lys	Leu	Glu	Pro	Gly	Gly	Lys	Туг		
					1990)				1995	5				2000)		
25	CAC	GTC	ATT	GTG	CAG	CTG	GGG	AAC	ATG	AGC	AAA	GAT	GCC	AGT	GTG	AAG	6228	
	His	Val	Ιle	Val	GIn	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ala	Ser	Val	Lys		
				2005	5				2010)				2015				
30	ATC	ACC	ACC	GTT	TCG	TTA	TCG	GCA	CCC	GAT	GCC	TTA	AAA	ATC	ATA	ACA	6276	
	Ile	Thr	Thr	Val	Ѕег	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	Ile	lle	Thr		
			2020)		2025				;			2030)				
35	GAA	AAT	GAC	CAC	GTC	CTT	CTC	TTC	TGG	AAA	AGT	CTA	GCT	СТА	AAG	GAA	6324	
	0.1																	
	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys	Glu		
40	Glu	Asn 2035		His	Val		Leu 2040		Trp	Lys	Ser	Leu 2045		Leu	Lys	Glu		
40		2035	ō		(2040					2045	i				6372	
	AAG	2035 TAT	ī 171	AAC	c GAA	AGC	2040 agg)	TAC	GAG	АТА	2045 CAC	ATG	ፐፐፕ	GAΤ	AGC	6372	
40 45	AAG	2035 TAT Tyr	ī 171	AAC	c GAA	AGC	2040 AGG Arg	GGC	TAC	GAG	АТА	2045 CAC His	ATG	ፐፐፕ	GAΤ	AGC	6372	
·	AAG Lys 2050	2035 TAT Tyr)	TTT Phe	AAC Asn	GAA Glu	AGC Ser 2055	2040 AGG Arg	GGC	TAC Tyr	GAG Glu	ATA 11e 2060	2045 CAC His	ATG Met	TTT Phe	GAT Asp	AGC Ser 2065	6372 6420	
45	AAG Lys 2050 GCC	2035 TAT Tyr) ATG	TTT Phe	AAC Asn ATC	GAA Glu	AGC Ser 2055 GCA	2040 AGG Arg	GGC Gly	TAC Tyr GGG	GAG Glu AAT	ATA Ile 2060	2045 CAC His	ATG Met	TTT Phe	GAT Asp	AGC Ser 2065 TTT		
45	AAG Lys 2050 GCC	2035 TAT Tyr) ATG	TTT Phe	AAC Asn ATC	GAA Glu	AGC Ser 2055 GCA Ala	2040 AGG Arg	GGC Gly	TAC Tyr GGG	GAG Glu AAT	ATA Ile 2060 ACT Thr	2045 CAC His	ATG Met	TTT Phe AAT Asn	GAT Asp	AGC Ser 2065 TTT Phe		

	Lys	lle	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Туг	Thr	Phe	Thr	Val	Gln	
5				2085	5				2090)				2095	5		
	GCA	CGA	TGC	CTT	TTG	GGC	AGC	CAG	ATC	TGC	GGG	GAG	CCT	GCC	GTG	CTA	6516
	Ala	Arg	Cys	Leu	Leu	Gly	Ser	Gln	Ιlе	Cys	Gly	Glu	Pro	Ala	Val	Leu	
10			2100)				2105	5		•		2110)			
	CTG	TAT	GAT	GAG	CTG	GGG	TCT	GGT	GGC	GAT	GCG	TCG	GCG	ATG	CAG	GCT	6564
15	Leu	Туг	Asp	Glu	Leu	Gly	Ser	Gly	Gly	Asp	Ala	Ser	Ala	Met	Gln	Ala	
, 5		2115	5				2120)				2125	5				
	GCC	AGG	TCT	ACT	GAT	GTC	GCC	GCC	GTG	GTG	GTG	CCC	ATC	CTG	TTT	CTG	6612
20	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	lle	Leu	Phe	Leu	
	2130)				2135	5				2140)				2145	
	ATA	CTG	CTG	AGC	CTG	GGG	GTC	GGG	TTT	GCC	ATC	CTG	TAC	ACG	AAG	CAT	6660
25	Ιlе	Leu	Leu	Ser	Leu	Gly	Val	Gly	Phe	Ala	Ile	Leu	Tyr	Thr	Lys	His	
					2150)				215	5						
	CGG	AGG	CTG	CAG	AGC	AGC	TTC	ACC	GCC	TTC	GCC	AAC	AGC	CAC	TAC	AGC	6708
30	Arg	Arg	Leu	Gln	Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	His	Tyr	Ser	
				216	5				217	0				217	5		
as	TCC	AGA	CTC	GGC	TCC	GCC	ATC	TTC	TCC	TCT	GGG	GAT	GAC	TTG	GGG	GAG	6756
35	Ser	Arg	Leu	Gly	Ser	Ala	lle	Phe	Ser	Ser	Gly	Asp	Asp	Leu	Gly	Glu	
			218	0				218	5				219	0			
40	GAT	GAT	GAA	GAT	GCT	CCT	ATG	ATC	ACT	GGA	TTT	TCG	GAC	GAC	GTC	CCC	6804
	Asp	Asp	Glu	Asp	Ala	Pro	Met	Ιle	Thr	Gly	Phe	Ser	Asp	Asp	Val	Pro	
		219	5				220	0				220	5				
45	ATG	GTG	АТА	GCC	TGA	AAGA	GCT	TTCC	TCAC	TA G	AAAC	CAAA	T GO	TGTA	AATA		6856
	Met	Val	Ile	Ala													
	221	0															
50	TTT	ТАТТ	TGA	TAAA	.GATA	GT T	GATO	GTTT	'A TT	TT AA	AAGA	TGC	CACTI	TGA	GTTC	CAATAT	6916
	GTT	'ATTT	ATT	TATO	GGCC	AA A	AACA	AAAG	C AA	AAAA	AAAA	AAA	AA				6961

	Sequence ID No. 4	
5	Length of the Sequence: 300	
	Type: nucleic acid	
	Strandedness: double	
10	Topology: linear	
	Molecular type: cDNA to mRNA	
	Sequence:	
15	ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCTGACC ATGGAGAGTG	60
	ATATCAAGGT GAATGGCTAT GTGGTGAACC TTTTCTGGGC ATTTGACACC CACAAGCAAG	120
20	AGAGGAGAAC TTTGAACTTC CGAGGAAGCA TATTGTCACA CAAAGTTGGC AATCTGACAG	180
	CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGGGAT AGCCCTCTGG	240
	CATTTGAGCA TGTTATGACC AGAGGGGTTC GCCCACCTGC ACCTAGCCTC AAGGCCAAAG	300
25	Sequence ID No. 5	
	Length of the Sequence: 6642	
	Type: nucleic acid	
30	Strandedness: double	
	Topology: linear	
<i>35</i>	Molecular type: cDNA to mRNA	
	Sequence:	
	ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCGT TCCTATTCAC CCTGGTCGCA	60
40	CTGCTGCCGC CCGGAGCTCT CTGCGAAGTC TGGACGCAGA GGCTGCACGG CGGCAGCGCG	120
	CCCTTGCCCC AGGACCGGGG CTTCCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG	180
	TGGGCGCGC GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CCGGAGGAAA	240
45	CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT	300
	TCCCACAATC AGATGGTGGT GCACTGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG	360
50	GCCCGAGATA GCCTGGCATT GGCGAGGCCC AAGAGCAGTG ATGTGTACGT GTCTTACGAC	420
	TATGGAAAAT CATTCAAGAA AATTTCAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT	480

	GAAGCTGTTA	TCGCCCAGTT	CTACCACAGC	CCTGCGGACA	ACAAGCGGTA	CATCTTTGCA	540
5	GACGCTTATG	CCCAGTACCT	CTGGATCACG	TTTGACTTCT	GCAACACTCT	TCAAGGCTTT	600
	TCCATCCCAT	TTCGGGCAGC	TGATCTCCTC	CTACACAGTA	AGGCCTCCAA	CCTTCTCTTG	660
	GGCTTTGACA	GGTCCCACCC	CAACAAGCAG	CTGTGGAAGT	CAGATGACTT	TGGCCAGACC	720
10	TGGATCATGA	TTCAGGAACA	TGTCAAGTCC	TTTTCTTGGG	GAATTGATCC	CTATGACAAA	780
	CCAAATACCA	TCTACATTGA	ACGACACGAA	CCCTCTGGCT	ACTCCACTGT	CTTCCGAAGT	840
	ACAGATTTCT	TCCAGTCCCG	GGAAAACCAG	GAAGTGATCC	TTGAGGAAGT	GAGAGATTTT	900
15	CAGCTTCGGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTGAACAG	960
	CAGTCTTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCGGAAGC	CCATGAGAGC	AGCCCAGTTT	1020
00	GTCACAAGAC	АТССТАТТАА	TGAATATTAC	ATCGCAGATG	CCTCCGAGGA	CCAGGTGTTT	1080
20	GTGTGTGTCA	GCCACAGTAA	CAACCGCACC	AATTTATACA	TCTCAGAGGC	AGAGGGGCTG	1140
	AAGTTCTCCC	TGTCCTTGGA	GAACGTGCTC	TATTACAGCC	CAGGAGGGC	CGGCAGTGAC	1200
25	ACCTTGGTGA	GGTATTTTGC	AAATGAACCA	TTTGCTGACT	TCCACCGAGT	GGAAGGATTG	1260
	CAAGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCG	1320
	GTCATCACCT	TTGACAAAGG	GGGAACCTGG	GAGTTTCTTC	AGGCTCCAGC	CTTCACGGGA	1380
30	TATGGAGAGA	AAATCAATTG	TGAGCTTTCC	CAGGGCTGTT	CCCTTCATCT	GGCTCAGCGC	1440
	CTCAGTCAGC	TCCTCAACCT	CCAGCTCCGG	AGAATGCCCA	TCCTGTCCAA	GGAGTCGGCT	1500
	CCAGGCCTCA	TCATCGCCAC	TGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
35	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	ТСАСТАСТАС	1620
	ACATGGGGAG	ACCACGGCGG	AATCATCACG	GCCATTGCCC	AGGGCATGGA	AACCAACGAG	1680
40	CTAAAATACA	GTACCAATGA	AGGGGAGACC	TGGAAAACAT	TCATCTTCTC	TGAGAAGCCA	1740
	GTGTTTGTGT	ATGGCCTCCT	CACAGAACCT	GGGGAGAAGA	GCACTGTCTT	CACCATCTTT	1800
	GGCTCGAACA	AAGAGAATGT	CCACAGCTGG	CTGATCCTCC	AGGTCAATGC	CACGGATGCC	1860
45	TTGGGAGTTC	CCTGCACAGA	GAATGACTAC	AAGCTGTGGT	CACCATCTGA	TGAGCGGGG	1920
	AATGAGTGTT	TGCTGGGACA	CAAGACTGTT	TTCAAACGGC	GGACCCCCCA	TGCCACATGC	1980
	TTCAATGGAG	AGGACTTTGA	CAGGCCGGTG	GTCGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
50	GACTATGAGT	GTGACTTCGG	TTTCAAGATG	AGTGAAGATT	TGTCATTAGA	GGTTTGTGTT	2100
	CCAGATCCGG	AATTTTCTGG	AAAGTCATAC	тесестеств	TGCCTTGCCC	TGTGGGTTCT	2160

	ACTTACAGGA	GAACGAGAGG	CTACCGGAAG	ATTTCTGGGG	ACACTTGTAG	CGGAGGAGAT	2220
	GTTGAAGCGC	GACTGGAAGG	AGAGCTGGTC	CCCTGTCCCC	TGGCAGAAGA	GAACGAGTTC	2280
5	ATTCTGTATG	CTGTGAGGAA	ATCCATCTAC	CGCTATGACC	TGGCCTCGGG	AGCCACCGAG	2340
	CAGTTGCCTC	TCACCGGGCT	ACGGGCAGCA	GTGGCCCTGG	ACTTTGACTA	TGAGCACAAC -	2400
10	TGTTTGTATT	GGTCCGACCT	GGCCTTGGAC	GTCATCCAGC	GCCTCTGTTT	GAATGGAAGC	2460
	ACAGGGCAAG	AGGTGATCAT	CAATTCTGGC	CTGGAGACAG	TAGAAGCTTT	GGCTTTTGAA	2520
	CCCCTCAGCC	AGCTGCTTTA	CTGGGTAGAT	GCAGGCTTCA	AAAAGATTGA	GGTAGCTAAT	2580
15	CCAGATGGCG	ACTTCCGACT	CACAATCGTC	AATTCCTCTG	TGCTTGATCG	TCCCAGGGCT	2640
	CTGGTCCTCG	TGCCCCAAGA	GGGGGTGATG	TTCTGGACAG	ACTGGGGAGA	CCTGAAGCCT	2700
	GGGATTTATC	GGAGCAATAT	GGATGGTTCT	GCTGCCTATC	ACCTGGTGTC	TGAGGATGTG	2760
20	AAGTGGCCCA	ATGGCATCTC	TGTGGACGAC	CAGTGGATTT	ACTGGACGGA	TGCCTACCTG	2820
	GAGTGCATAG	AGCGGATCAC	GTTCAGTGGC	CAGCAGCGCT	CTGTCATTCT	GGACAACCTC	2880
25	CCGCACCCCT	ATGCCATTGC	TGTCTTTAAG	AATGAAATCT	ACTGGGATGA	CTGGTCACAG	2940
	CTCAGCATAT	TCCGAGCTTC	CAAATACAGT	GGGTCCCAGA	TGGAGATTCT	GGCAAACCAG	3000
	CTCACGGGGC	TCATGGACAT	GAAGATTTTC	TACAAGGGGA	AGAACACTGG	AAGCAATGCC	3060
30	TGTGTGCCCA	GGCCATGCAG	CCTGCTGTGC	CTGCCCAAGG	CCAACAACAG	TAGAAGCTGC	3120
	AGGTGTCCAG	AGGATGTGTC	CAGCAGTGTG	CTTCCATCAG	GGGACCTGAT	GTGTGACTGC	3180
	CCTCAGGGCT	ATCAGCTCAA	GAACAATACC	TGTGTCAAAG	AAGAGAACAC	CTGTCTTCGC	3240
35	AACCAGTATC	GCTGCAGCAA	CGGGAACTGT	ATCAACAGCA	TTTGGTGGTG	TGACTTTGAC	3300
	AACGACTGTG	GAGACATGAG	CGATGAGAGA	AACTGCCCTA	CCACCATCTO	TGACCTGGAC	3360
40	ACCCAGTTTC	GTTGCCAGGA	GTCTGGGACT	TGTATCCCAC	TGTCCTATAA	ATGTGACCTT	3420
40	GAGGATGACT	GTGGAGACAA	CAGTGATGAA	AGTCATTGTG	AAATGCACCA	GTGCCGGAGT	3480
	GACGAGTACA	ACTGCAGTTC	CGGCATGTGC	ATCCGCTCCT	CCTGGGTATO	TGACGGGGAC	3540
45	AACGACTGCA	GGGACTGGTC	TGATGAAGCO	AACTGTACCG	CCATCTATCA	CACCTGTGAG	3600
	GCCTCCAACT	TCCAGTGCCC	AAACGGCAC	TGCATCCCCC	AGCGGTGGG	GTGTGACGGG	3660
	GATACGGACT	GCCAGGATGC	TTCCGATGA(GATCCAGTCA	ACTGTGAGA	GAAGTGCAAT	3720
50	GGATTCCGCT	GCCCAAACGC	CACTTGCAT	CCATCÇAGCA	AACATTGTG/	TGGTCTGCGT	3780
	GATTGCTCTC	ATGGCTCCGA	TGAACAGCAG	TGCGAGCCC	TCTGTACGC	A CTTCATGGAC	3840

	TTTGTGTGTA	AGAACCGCCA	GCAGTGCCTG	TTCCACTCCA	TGGTCTGTGA	CGGAATCATC	3900
5	CAGTGCCGCG	ACGGGTCCGA	TGAGGATGCG	GCGTTTGCAG	GATGCTCCCA	AGATCCTGAG	3960
	TTCCACAAGG	TATGTGATGA	GTTCGGTTTC	CAGTGTCAGA	ATGGAGTGTG	CATCAGTTTG	4020
	ATTTGGAAGT	GCGACGGGAT	GGATGATTGC	GGCGATTATT	CTGATGAAGC	CAACTGCGAA	4080
10	AACCCCACAG	AAGCCCCAAA	CTGCTCCCGC	TACTTCCAGT	TTCGGTGTGA	GAATGGCCAC	4140
	TGCATCCCCA	ACAGATGGAA	ATGTGACAGG	GAGAACGACT	GTGGGGACTG	GTCTGATGAG	4200
	AAGGATTGTG	GAGATTCACA	TATTCTTCCC	TTCTCGACTC	CTGGGCCCTC	CACGTGTCTG	4260
15	CCCAATTACT	ACCGCTGCAG	CAGTGGGACC	TGCGTGATGG	ACACCTGGGT	GTGCGACGGG	4320
	TACCGAGATT	GTGCAGATGG	CTCTGACGAG	GAAGCCTGCC	CCTTGCTTGC	AAACGTCACT	4380
20	GCTGCCTCCA	CTCCCACCCA	ACTTGGGCGA	TGTGACCGAT	TTGAGTTCGA	ATGCCACCAA	4440
	CCGAAGACGT	GTATTCCCAA	CTGGAAGCGC	TGTGACGGCC	ACCAAGATTG	CCAGGATGGC	4500
	CGGGACGAGG	CCAATTGCCC	CACACACAGC	ACCTTGACTT	GCATGAGCAG	GGAGTTCCAG	4560
25	TGCGAGGACG	GGGAGGCCTG	CATTGTGCTC	TCGGAGCGCT	GCGACGGCTT	CCTGGACTGC	4620
	TCGGACGAGA	GCGATGAAAA	GGCCTGCAGT	GATGAGTTGA	CTGTGTACAA	AGTACAGAAT	4680
	CTTCAGTGGA	CAGCTGACTT	CTCTGGGGAT	GTGACTTTGA	CCTGGATGAG	GCCCAAAAA	4740
30	ATGCCCTCTG	CATCTTGTGT	ATATAATGTC	TACTACAGGG	TGGTTGGAGA	GAGCATATGG	4800
	AAGACTCTGG	AGACCCACAG	CAATAAGACA	AACACTGTAT	TAAAAGTCTT	GAAACCAGAT	4860
	ACCACGTATC	AGGTTAAAGT	ACAGGTTCAG	TGTCTCAGCA	AGGCACACAA	CACCAATGAC	4920
35	TTTGTGACCC	TGAGGACCCC	AGAGGGATTG	CCAGATGCCC	CTCGAAATCT	CCAGCTGTCA	4980
	CTCCCCAGGG	AAGCAGAAGG	TGTGATTGTA	GGCCACTGGG	CTCCTCCCAT	CCACACCCAT	5040
40	GGCCTCATCC	GTGAGTACAT	TGTAGAATAC	AGCAGGAGTG	GTTCCAAGAT	GTGGGCCTCC	5100
	CAGAGGGCTG	CTAGTAACTT	TACAGAAATC	AAGAACTTAT	TGGTCAACAC	TCTATACACC	5160
	GTCAGAGTGG	CTGCGGTGAC	TAGTCGTGGA	ATAGGAAACT	GGAGCGATTC	TAAATCCATT	5220
45	ACCACCATAA	AAGGAAAAGT	GATCCCACCA	CCAGATATCC	ACATTGACAG	CTATGGTGAA	5280
	AATTATCTAA	GCTTCACCCT	GACCATGGAG	AGTGATATCA	AGGTGAATGG	CTATGTGGTG	5340
	AACCTTTTCT	GGGCATTTGA	CACCCACAAG	CAAGAGAGGA	GAACTTTGAA	CTTCCGAGGA	5400
50	AGCATATTGT	CACACAAAGT	TGGCAATCTG	ACAGCTCATA	CATCCTATGA	GATTTCTGCC	5460
	TGGGCCAAGA	CTGACTTGGG	GGATAGCCCT	CTGGCATTTG	AGCATGTTAT	GACCAGAGGG	5520

	GTTCGCCCAC CTGCACCTAG CCTCAAGGCC AAAGCCATCA ACCAGACTGC AGTGGAATGT	5580
5	ACCTGGACCG GCCCCCGGAA TGTGGTTTAT GGTATTTTCT ATGCCACGTC CTTTCTTGAC	5640
	CTCTATCGCA ACCCGAAGAG CTTGACTACT TCACTCCACA ACAAGACGGT CATTGTCAGT	5700
	AAGGATGAGC AGTATTTGTT TCTGGTCCGT GTAGTGGTAC CCTACCAGGG GCCATCCTCT	5760
10	GACTACGTTG TAGTGAAGAT GATCCCGGAC AGCAGGCTTC CACCCCGTCA CCTGCATGTG	5820
	GTTCATACGG GCAAAACCTC CGTGGTCATC AAGTGGGAAT CACCGTATGA CTCTCCTGAC	5880
	CAGGACTTGT TGTATGCAAT TGCAGTCAAA GATCTCATAA GAAAGACTGA CAGGAGCTAC	5940
15	AAAGTAAAAT CCCGTAACAG CACTGTGGAA TACACCCTTA ACAAGTTGGA GCCTGGCGGG	6000
	AAATACCACA TCATTGTCCA ACTGGGGAAC ATGAGCAAAG ATTCCAGCAT AAAAATTACC	6060
20	ACAGTTTCAT TATCAGCACC TGATGCCTTA AAAATCATAA CAGAAAATGA TCATGTTCTT	6120
20	CTCTTTTGGA AAAGCCTGGC TTTAAAGGAA AAGCATTTTA ATGAAAGCAG GGGCTATGAG	6180
	ATACACATGT TTGATAGTGC CATGAATATC ACAGCTTACC TTGGGAATAC TACTGACAAT	6240
25	TTCTTTAAAA TTTCCAACCT GAAGATGGGT CATAATTACA CGTTCACCGT CCAAGCAAGA	6300
	TGCCTTTTTG GCAACCAGAT CTGTGGGGAG CCTGCCATCC TGCTGTACGA TGAGCTGGGG	6360
	TCTGGTGCAG ATGCATCTGC AACGCAGGCT GCCAGATCTA CGGATGTTGC TGCTGTGGTG	6420
30	GTGCCCATCT TATTCCTGAT ACTGCTGAGC CTGGGGGTGG GGTTTGCCAT CCTGTACACG	6480
	AAGCACCGGA GGCTGCAGAG CAGCTTCACC GCCTTCGCCA ACAGCCACTA CAGCTCCAGG	6540
	CTGGGGTCCG CAATCTTCTC CTCTGGGGAT GACCTGGGGG AAGATGATGA AGATGCCCCT	6600
35	ATGATAACTG GATTTTCAGA TGACGTCCCC ATGGTGATAG CC	6642
	Sequence ID No. 6	
40	Length of the Sequence: 2214	
	Type: amino acid	
	Topology: linear	
45	Molecular type: Protein	
	Sequence:	
	Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe	
50	5 10 15	
	Thr Ley Val Ala Ley Ley Pro Pro Gly Ala Ley Cys Gly Val Trp Thr	

					20)				25	5				30)	
5	G	ln i	Arg	Leu	His	Gly	Gly	/ Ser	Ala	Pro) Leu	Pro	Gln	Asp	Arg	Gly	Phe
				35					40					45			
	L	eu 1	Val	Val	Gln	G1 y	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Тгр	Ala	Arg	Gly
10			50					55					60				
	A	sp A	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	Lys	Pro	Leu	Arg	Arg	Lys
15		35					70					7 5					80
15	A	-g 5	Зег	Ala	Ala	Leu	Gln	Pro	Glu	Pro	lle	Lys	Val	Туг	Gly	Gln	Val
						85					90					95	
20	Se	er L	.eu	Asn	Asp	Ser	His	Asn	Gln	Met	Val	Val	His	Trp	Ala	Gly	Glu
					100					105					110		
	Ly	's S	er	Asn	Val	Ile	Val	Ala	Leu	Ala	Arg	Asp	Ser	Leu	Ala	Leu	Ala
25				115					120					125			
	Aı	g F	o 1°	Lys	Ser	Ser	Asp	Va]	Туг	Val	Ser	Tyr	Asp	Туг	Gly	Lys	Ser
		1	30					135					140				
30	Pł	e L	ys	Lys	Ιlе	Ser	Asp	Lys	Leu	Asn	Phe	Gly	Leu	Gly	Asn	Arg	Ser
	14	5					150					155					160
35	Gl	u A	la	Val	Ile	Ala	Gln	Phe	Туг	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg
33						165					170					175	
	Ту	r I	le	Phe	Ala	Asp	Ala	Туг	Ala	Gln	Tyr	Leu	Trp	Ιlе	Thr	Phe	Asp
40					180					185					190		
	Ph	e C	ys	Asn	Thr	Leu	Gln	Gly	Phe	Ser	Ile	Pro	Phe	Arg	Ala	Ala	Asp
				195					200					205			
45	Le	u L	eu	Leu	His	Ser	Lys	Ala	Ser	Asn	Leu	Leu	Leu	Gly	Phe	Asp	Arg
		2	10					215					220				
	Se	r H	is	Pro	Asn	Lys	Gln	Leu	Тгр	Lys	Ser	Asp	Asp	Phe	Gly	Gln	Thr
50	22	5					230					235					240
	Tr	p II	le l	Met	lle	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Тгр	Gly	He	Asp

					245					250					255	
5	Pro	Tyr	Asp	Lys	Pro	Asn	Thr	He	Туг	He	Glu	Arg	His	Glu	Pro	Ser
				260		•			265					270		
	Gly	Туr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gln	Ser	Arg	Glu
10			275					280					285			
	Asn	Gln	Glu	Val	He	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp
		290					295					300				
15	Lys	Туг	Met	Phe	Ala	Thr	Lys	Val	Val	His	Leu	Leu	Gly	Ser	Glu	Gln
	305					310					315					320
	Gln	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg
20					325					330					335	
	Ala	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	He	Asn	Glu	Tyr	Туг	lle	Ala
25				340					345					350		
	Asp	Ala	Ser	Glu	Asp	Gln	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn
	-		355					360					365			
30	Arg	Thr		Leu	Туг	He	Ser	Glu	Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu
		370					375					380				
	Ser	Leu	Glu	Asn	Val	Leu	Tyr	Tyr	Ser	Pro	Gly	Gly	Ala	Gly	Ser	Asp
35	385					390					395					400
			Va!	Arg	Туг	Phe	Ala	Asn	Glu	Pro	Phe	Ala	Asp	Phe	His	Arg
					405					410					415	
40	Val	Glu	Gly	Leu	Gln	Gly	Val	Туг	He	Ala	Thr	Leu	lle	Asn	Gly	Ser
				420					425					430		
45	Met	Asn	Glu	Glu	Asn	Met	Arg	Ser	Val	He	Thr	Phe	Asp	Lys	Gly	Gly
			435					440					445			
	Thr	Trp	Glu	Phe	Leu	Gln	Ala	Pro	Ala	Phe	Thr	Gly	Tyr	Gly	Glu	Lys
50		450					455					460				
	He			Glu	Leu	Ser	Gln	Gly	Cys	Ser	Leu	His	Leu	Ala	Gin	Arg

	465					470					475					480
5	Leu	Ser	Gln	Leu	Leu	Asn	Leu	Gln	Leu	Arg	Arg	Met	Pro	Ιle	Leu	Ser
•					485					490					495	
	Lys	G1 u	Ser	Ala	Pro	Gly	Leu	Ιlе	Ile	Ala	Thr	Gly	Ser	Val	Gly	Lys
10				500					505					510		
	Asn	Leu	Ala	Ser	Lys	Thr	Asn	Val	Tyr	Ile	Ser	Ser	Ser	Ala	Gly	Ala
			515					520					525			
15	Arg	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Tyr	Tyr	Thr	Trp	Gly	Asp
		530					535					540				
	His	Gly	Gly	He	lle	Thr	Ala	Ιle	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu
20	545					550					555					560
	Leu	Lys	Tyr	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	Thr	Phe	Ιlе	Phe
25					565					570					575	
	Ser	Glu	Lys	Pro	Val	Phe	Val	Tyr	Gly	Leu	Leu	Thr	Glu	Pro	Gly	Glu
				580					585				-	590		
30	Lys	Ser	Thr	Val	Phe	Thr	lle	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His
			595					600					605			
	Ser	Trp	Leu	Пе	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Pro
35		610					615					620				
	Cys	Thr	Glu	Asn	Asp	Туг	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly
42	625					630					635					640
10	Asn	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro
					645					650					655	
1 5	His	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val
				660					665					670		
	Ser	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Туг	Glu	Cys	Asp	Phe	Gly	Phe
50			675					680					685			
	Lys	Met	Ser	Glu	Asp	Leu	Ser	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Glu

		690					695					700				
	Phe	Ser	Gly	Lys	Ser	Туг	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Ser
5	705					710					715					720
,	Thr	Туг	Arg	Arg	Thr	Arg	Gly	Tyr	Arg	Lys	Ile	Ser	Gly	Asp	Thr	Cys
10					725					730					735	
	Ser	Gly	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Cys
				740					745					7 50		
15	Pro	Leu	Ala	Glu	Glu	Asn	Glu	Phe	He	Leu	Tyr	Ala	Val	Arg	Lys	Ser
			755					760					765			
	I l e	Tyr			Asp	Leu	Ala	Ser	Gly	Ala	Thr	Glu	Gln	Leu	Pro	Leu
20		770					775					780				
	Thr			Arg	Ala	Ala	. Val	Ala	Leu	Asp	Phe	Asp	Tyr	Glu	His	Asn
	785					790					795					800
25			туг	Trp	Ser	- Ası	Lei	ı Ala	a Lei	ı Ası	vàl	lle	Glr	Arg	Lei	Cys
	•				809					810					815	
30	Lei	ı Ası	ı Gly	y Sei	r T hi	r Gl	y Gli	n Gli	ı Va	1 11	e ile	e Asr	\ Sei	Gly	/ Lei	ı Glu
				820					82					830		
	Th	r Va	1 G1:			u Al	a Ph	e Gl	u Pr	o Le	u Se	r Gli	n Lei	u Le	з Ту	r Trp
35			83					84					84			
	۷a	ıl As			y Ph	ie Ly	s Ly	s II	e Gl	u Va	ıl Al	a As	n Pr	o As	p Gl	y Asp
		85					85					86				
40	PI			eu Th	nr II	le Va	al As	sn Se	er Se	er Va	al Le	eu As	p Ar	g Pr	o Ar	g Ala
•		35	•				70					75				880
			al Lo	eu Va	al Pi			lu G	ly Va	al M	et Pl	ne Tr	p Tł	nr As	sp Ті	p Gly
45	~	· ·				85					90					95
	Δ	sn L	en L	vs P			le T	уг А	rg S	er A	sn M	et As	sp G	ly S	er A	la Ala
50	. ^	Jp J			00	•				05					10	
J .	π	vr H	is L			er G	ilu A	sp V			rp P	ro A	sn G	ly I	le S	er Val
		y		J _ '				-								

			915	<u>, </u>				920)				92	5		
5	Asp	Asp	Gln	Tr p	Ile	э Туг	Tr	p Thi	As	p Al	а Ту	r Le	u Gl	и Су:	s II	e Glu
		930					939					94				
	Arg	He	Thr	Phe	Ser	Gly	/ Gli	n Glr	Arg	g Sei	r Val	11	e Le	u Asp) Ası	n Leu
10	945					950					955					960
	Pro	His	Pro	Туг	Ala	Пε	Ala	ı Val	Phe	e Lys	s Asr	GI	u Ile	e Tyr	Tr) Asp
					965					970					975	
15	Asp	Trp	Ser	Gln	Leu	Ser	He	Phe	Arg	Ala	a Ser	Ly:	s Tyı	Ser	Gly	/ Ser
				980					985					990		
	Gln	Met	Glu	Пe	Leu	Ala	Asn	Gln	Leu	Thr	Gly	Lei	ı Met	: Asp	Met	Lys
20			995					100	0				100)5 .		
	Ile	Phe	Tyr	Lys	Gly	Lys	Asn	Thr	Gly	Ser	Asn	Ala	a Cys	Val	Pro	Arg
25		1010					101					102				
	Pro	Cys	Ser	Leu	Leu	Cys	Leu	Pro	Lys	Ala	Asn	Asr	ı Ser	Arg	Ser	Cys
	1025					103					103					1040
30	Arg	Cys	Pro	Glu	Asp	Val	Ser	Ser	Ser	Val	Leu	Pro	Ser	Gly	Asp	
					1045					105					105	
	Met	Cys	Asp	Cys	Pro	Gln	Gly	Туг	Gln	Leu	Lys	Asn	Asn	Thr		
35				1060					106					107		
	Lys (Glu (Glu <i>i</i>	Asn ′	lhr	Cys	Leu	Arg	Asn	Gln	Tyr	Arg	Cys	Ser	Asn	Gly
40			1075					1080					1085			•
40	Asn C	ys I	lle A	lsn S	Ser	lle	Тгр	Trp	Cys	Asp	Phe	Asp	Asn	Asp	Cys	Gly
		090					1095					1100				•
							A = -	C	D== :	Φ L	ТЬ -	Па	Cuo			A c n
45	Asp M	let S	Ser A	usp (ilu <i>l</i>	Arg .	ASN	Cys .	710	ınr	1111	116	Cys	Asp	Leu	42h
45	Asp M 1105	let S	Ser A	sp C		arg . 1110	ASN	Cys .	riu		1115		Cys	Asp	Leu	1120
45	1105				1	1110					1115					1120
50				rg C	1	1110			Gly '		1115			Leu	Ser	1120 Т уг
	1105	ln P	he A	irg C	ys (125	1110 Gln (Glu :	Ser (Gly '	Thr 1130	1115 Cys	lle	Pro	Leu	Ser 1135	1120 Tyr

				1140)				1145	j				1150)		
	Cys	Glu	Met	His	Gln	Cys	Arg	Ser	Asp	Glu	Tyr	Asn	Cys	Ser	Ser	Gl	У
5			115	5				1160)				1165	5			
	Met	Cys	He	Arg	Ser	Ser	Trp	Val	Cys	Asp	Gly	Asp	Asn	Asp	Cys	Ar	g
10		117	0				1175	5				1180)				
	Asp	Trp	Ser	Asp	Glu	Ala	Asn	Cys	Thr	Ala	He	Tyr	His	Thr	Cys	Gl	u
	118	5				119	0				119	5				12	200
15	Ala	Ser	Asn	Phe	Gln	Cys	Arg	Asn	Gly	His	Cys	Ιlе	Pro	Gln	Arg	Tr	. b
					120	5				121)				121	.5	
	Ala	Cys	Asp	Gly	Asp	Thr	Asp	Cys	G1 n	Asp	Gly	Ser	Asp	Glu	Asp	Pı	ro
20				122	:0				122	5				123	0		
	Val	Asn	Cys	Glu	Lys	Lys	Cys	Asn	Gly	Phe	Arg	Cys	Pro	Asr	Gly	/ T	hr
25			123	35				124	.0				124	5			
	Cys	s Ile	e Pro	Ser	Ser	Lys	His	Cys	Asp	Gly	Leu	Arg	Asp	Cys	s Se	r A	sp
		125	50				125	55				126	60				
30	Gl	y Sei	r Ası	p Glu	ı Glı	His	s Cys	s Glu	ı Pro	Leu	Cys	Thi	His	s Ph	e Me	t A	.sp
	12	65				12	70				12"	75				1	280
	Ph	e Va	l Cy	s Ly	s As	n Ar	g Gl	n Gli	n Cy:	s Lei	ı Phe	e His	s Se	r Me	t Va	1 0	:ys
35					12	85		•		129	90				12	95	
	As	p Gl	y Il	e II	e Gl	n Cy	s Ar	g As	p Gl	y Se	r As	p Gl	u As	p Al	a Al	a I	?he
40				. 13	300				13	805				13	310		
40	Αl	a Gl	y Cy	ıs Se	er Gl	n As	sp Pr	o Gl	u Pł	ne Hi	s Ly	's Va	ıl Cy	s As	sp G	l u	Phe
			13	315				13	320				13	325			
45	G!	y Ph	ne Gl	ln Cy	s Gl	n As	sn Gl	y Va	ıl Cy	s Il	e Se	r Le	eu II	e Ti	p L	ys	Cys
		13	330				13	335				13	340				
	As	sp G	ly Me	et As	sp As	sp Cy	/s Gl	y As	зр Ту	yr Se	er As	sp G1	u Al	a As	sn C	ys	Glu
50	13	345				13	350				13	355					1360
	A:	sn Pi	ro Ti	hr G	lu A	la Pi	ro As	sn Cy	/s S	er Ai	g Ty	r Pl	ne G	in Pi	he A	rg	Cys

				136	35				137	70				13	75
<i>E</i>	Glu	Asn G	ly Hi	s Cys	s II	e Pr	o As	n Arg	g Trp	Lys	S Cy:	s As	p Ar	g Gli	ı Asn
5			13					138					139		
	Asp	Cys G	ly As _i	p Trp	Se	r As	p Gl	u Lys	s Asp	Cys	Gly	y As	p Sei	· His	lle
10			395				- 14					14			
	Leu	Pro Ph	ie Sei	Thr	Pre	o Gly	y Pro	o Ser	Thr	Cys	Leu	Pr	o Asr	Туг	Туг
		1410				14					142				•
15	Arg (Cys Se	r Ser	Gly	Thr	Cys	s Val	Met	Asp	Thr	Тгр	Va:	l Cys	Asp	Glv
	1425				143					143			·	•	1440
	Туг А	rg As	p Cys	Ala	Asp	Gly	Ser	Asp	Glu	Glu	Ala	Cys	s Pro	Leu	
20				144					1450					145	
	Ala A	sn Va	l Thr	Ala	Ala	Ser	Thr	Pro	Thr	Gln	Leu	Gly	Arg		
25			146					1465					1470		
	Arg P	he Gli	u Phe	Glu	Cys	His	Gln	Pro	Lys	Thr	Cys	He			Tro
		14'					148					148			p
30	Lys A	rg Cys	s Asp	Gly	His	Gln	Asp	Cys	Gln	Asp	Glv			Gla	Ala
		490				149					1500				
	Asn C	s Pro	Thr	His	Ser	Thr	Leu	Thr	Cys	Met			Glu	Phe	Gln
35	1505				1510					1 5 15					1520
	Cys G	u Asp	Gly	Glu	Ala	Cys	He	Val				Arg	Cvs	Asn	
40				1525					1530			J		1535	
40	Phe Le	u Asp	Cys	Ser /	Asp	Glu	Ser	Asp (Glu i	Lvs	Ala	Cvs	Ser		
			1540					1545		_,		0,0	1550		o i u
45	Leu Th	r Val	Tyr	Lys \	/al	Gln			Gln 7	îrp 1	Chr ,	Ala			Ser
		155					1560					1565			JC1
	Gly As	p Val	Thr	Leu 1	`hr'				oro L	vs I				Ser A	ıla
50	15					1575		•			580			,,,,	114
	Ser Cy:	s Val	Tyr /	Asn V			ſyr <i>i</i>	Arg V	'al V			:1 11	Ser 1	ו מו	`rn
						-			'	٠. U			JC1 1	16 1	ıμ

	1589	5				1590)				1595	5				1600
	Lys	Thr	Leu	Glu	Thr	His	Ser	Asn	Lys	Thr	Asn	Thr	Val	Leu	Lys	Val
					1605	5				1610)				1615	5
	Leu	Lys	Pro	Asp	Thr	Thr	Туг	Gln	Val	Lys	Val	Gln	Val	Gln	Cys	Leu
10				1620)				1625	5				1630)	
	Ser	Lys	Ala	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu
			1635	5				1640)				1645	5		
15	Gly	Leu	Pro	Asp	Ala	Pro	Arg	Asn	Leu	Gln	Leu	Ser	Leu	Pro	Arg	Glu
		1650)				1655	5				1660)			
	Ala	Glu	Gly	Val	Ile	Val	Gly	His	Trp	Ala	Pro	Pro	He	His	Thr	His
20	1669	5				1670)				1675	5				1680
	Gly	Leu	Ile	Arg	Glu	Tyr	Ile	Val	Glu	Tyr	Ser	Arg	Ser	Gly	Ser	Lys
0 5					1685	5				1690)				1695	5
25	Met	Trp	Ala	Ser	Gln	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	Ιle	Lys	Asn
				1700)				1705	ō				1710)	
30	Leu	Leu	Val	Asn	Thr	Leu	Tyr	Thr	Val	Arg	Val	Ala	Ala	Val	Thr	Ser
			1715	5				1720)				1725	5		
	Arg	Gly	Ile	Gly	Asn	Trp	Ser	Asp	Ser	Lys	Ser	He	Thr	Thr	lle	Lys
35		1730)				1735	5				1740)			
	Gly	Lys	Val	lle	Pro	Pro	Pro	Asp	Ile	His	He	Asp	Ser	Туг	Gly	Glu
	1745	5				1750)				1759	5				1760
40	Asn	Tyr	Leu	Ser	Phe	Thr	Leu	Thr	Met	Glu	Ser	Asp	He	Lys	Val	Asn
					176	5				177	0				177	5
45	Gly	Tyr	Val	Val	Asn	Leu	Phe	Trp	Ala	Phe	Asp	Thr	His	Lys	Gln	Glu
45				1780)				178	5				179	0	
	Arg	Arg	Thr	Leu	Asn	Phe	Arg	Gly	Ser	He	Leu	Ser	His	Lys	Val	Gly
50			1795	5				1800)				180	5		
50	Asn	Leu			His	Thr	Ser			Ile	Ser	Ala			Lys	Thr

		18	10				18	15				18	20			
5	Asp	Le	u Gl	y As	p Se	r Pro	o Le	u Al	a Ph	e Glu	ı His	s Va	l Me	t Th	r Ar	gGly
	182					183					183					1840
	Val	Ar	g Pro	o Pr	o Ala	a Pro	Se	r Le	ц Ly	s Ala	a Lys	. Ala	a Ile	e Ası	n Glr	Thr
10					184					185					185	
	Ala	\ Va	l Glu	ı Cys	s Thr	Trp	Th	r Gl	y Pro	o Arg	, Asn	Val	l Val	Туг	Gly	lle
				186					186					187		
15	Phe	Туг	Ala	thr	Ser	Phe	: Lei	ı Ası	Lei	ı Tyr	Arg	Asr	Pro	Lys	Ser	Leu
			187					188					188			
20	Thr	Thr	Ser	Leu	His	Asn	Lys	Thr	Val	lle	Val	Ser	Lys	Asp	Glu	Gln
20		189					189					190				
	Tyr	Leu	Phe	Leu	Val	Arg	Val	Val	Val	Pro	Туг	Gln	Gly	Pro	Ser	Ser
25	1909					191					191					1920
	Asp	Tyr	Val	Val	Val	Lys	Met	Ιle	Pro	Asp	Ser	Arg	Leu	Рго	Pro	Arg
					192					193					193	
30	His	Leu	His	Val	Val	His	Thr	G1 y	Lys	Thr	Ser	Va!	Val	Ile	Lys	Тгр
				194					194					195		
	Glu	Ser	Pro	Tyr	Asp	Ser	Pro	Asp	Gln	Asp	Leu	Leu	Туг	Ala	lle	Ala
35			195					196			. •		1969			
	Val	Lys	Asp	Leu	Ile	Arg	Lys	Thr	Asp	Arg	Ser	Туг	Lys	Val	Lys	Ser
10		1970					1979					1980				
	Arg	Asn	Ser	Thr	Val	Glu	Tyr	Thr	Leu	Asn	Lys	Leu	Glu	Pro	Gly	Gly
	1985					1990					1995					2000
15	Lys '	Туг	His	He	He	Val	Gln	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ser	Ser
					2005					2010					2015	
	[le l	Lys	Ile	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro .	Asp	Ala			
o				2020					2025			•		2030		
	lle 1	ľhr	Glu	Asn	Asp i	His '	Val	Leu	Leu	Phe '	Trp i	Lys .				Leu
											•	• -				

			203	5				2040)				2049	5		
_	Lys	Glu	Lys	His	Phe	Asn	Glu	Ser	Arg	Gly	Туг	Glu	Пе	His	Met	Phe
5		2050)				2055	ō				2060)			
	Asp	Ser	Ala	Met	Asn	Ile	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn
10	2069	5				2070)				2075	5				2080
	Phe	Phe	Lys	Ile	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Phe	Thr
					2085	5				2090)				2095	5
15	Val	Gln	Ala	Arg	Cys	Leu	Phe	Gly	Asn	Gln	He	Cys	Gly	Glu	Pro	Ala
				2100)				2109	5				2110)	
	Ile	Leu	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Gly	Ala	Asp	Ala	Ser	Ala	Thr
2 0			2115	5				2120)				2125	5		
	Gln	Ala	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	Ile	Leu
25		2130)				2135	5				2140)			
25	Phe	Leu	lle	Leu	Leu	Ser	Leu	Gly	Val	Gly	Phe	Ala	He	Leu	Туг	Thr
	2145	5				2150)				2155	5				2160
30	Lys	His	Arg	Arg	Leu	Gln	Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	His
					2165	5				2170)				2175	5
•	Tyr	Ser	Ser	Arg	Leu	Gly	Ser	Ala	lle	Phe	Ser	Ser	Gly	Asp	Asp	Leu
35				2180)				2185	5				2190)	
	Gly	Glu	Asp	Asp	Glu	Asp	Ala	Pro	Met	He	Thr	Gly	Phe	Ser	Asp	Asp
			2195	5				2200)				2205	5		
40	Val	Pro	Met	Val	Ιlе	Ala										
		2210)													
45	Sequ	uenc	e I	D No	o. 7	1										
	Leng	gth	of	the	Seq	uen	ce:	684	.3							
~	Туре	∋: n	ucl	eic	aci	đ										
50	Stra	ande	dne	ss:	dou	ble										
	Topo	_														

	Mole	ecul	lar	typ	e: c	DNA	to	mRN	Α								
_	Feat	ture	e:														
5	Nar	ne/K	œy:	sig	g pe	pti	de										
	Loc	cati	on:	81	16	4											
10	Ide	enti	fic	atio	on π	eth	od:	s									
	Nan	ne/K	ey:	mat	t pe	pti	de										
	Loc	cati	on:	165	56	722											
15	Ide	enti	fic	atio	on m	eth	od:	s									
٠	Sequ	ienc	:e:														
											CCG	GCCC	CAGCO	GC 1	CTCC	TGGCC	23
20	TCGC	CGCTC	GCA (CATTO	CTCTO	CC TC	GGCGC	GCGGC	GCC	CACCI	rgca	GTAC	CGT1	rcg c	CCGA	ACATG	83
																Met	
25																1	
	GCG	ACA	CGG	AGC	AGC	AGG	AGG	GAG	TCG	CGA	CTC	CCG	TTC	CTA	TTC	ACC-	131
	Ala	Thr	Arg	Ser	Ser	Arg	Arg	Glu	Ser	Arg	Leu	Pro	Phe	Leu	Phe	Thr	
30				5					10					15			
					CTG												179
	Leu	Val	Ala	Leu	Leu	Pro	Pro	Gly	Ala	Leu	Cys	Glu	Val	Trp	Thr	Gln	
35			20					25					30				
					GGC												227
40	Arg		His	Gly	Gly	Ser		Pro	Leu	Pro	Gln		Arg	Gly	Phe	Leu	
		35					40					45					
					GAC												275
45		Val	GIn	Gly	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Trp	Ala	Arg	Gly	Asp	
	50					55					60					65	
					AGC												323
50	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	Lys	Pro	Leu	Arg	Arg	Lys	Arg	
					70					75					80		

	AGC	GCT	GCC	CTO	CAC	ccc	GAG	CC	CAT	C AA	G GT	G TAC	GGA	CAG	GTT	AGT	37
5	Ser	Ala	Ala	Leu	Glr	n Pro	Glu	Pro	o II	e Ly:	s Va	l Tyr	Gly	Gln	Val	Ser	
				85	,				90	0				95			
	CTG	AAT	GAT	TCC	CAC	СААТ	CAG	AT(G GTO	G GTO	G CA	C TGG	GCT	GGA	GAG	AAA	419
10	Leu	Asn	Asp	Ser	His	. Asn	Gln	Met	: Val	l Val	His	s Trp	Ala	Gly	Ğlu	Lys	
			100					105	5				110				
	AGC	AAC	GTG	ATC	GTG	GCC	TTG	GCC	C CG/	A GAT	r AG0	C CTG	GCA	TTG	GCG	AGG	467
15	Ser	Asn	Val	He	Val	Ala	Leu	Ala	L Arg	g Asp	Ser	Leu	Ala	Leu	Ala	Arg	
		115					120					125					
20	CCC	AAG	AGC	AGT	GAT	GTG	TAC	GTO	тст	TAC	GAC	тат	GGA	AAA	TCA	TTC	515
20	Pro	Lys	Ser	Ser	Asp	Val	Tyr	Val	Ser	Tyr	Asp	Tyr	Gly	Lys	Ser	Phe	
	130					135					140)				145	
25	AAG	AAA	ATT	TCA	GAC	AAG	TTA	AAC	111	GGC	TTC	GGA	AAT	AGG	AGT	GAA	563
	Lys	Lys	lle	Ser	Asp	Lys	Leu	Asn	Phe	Gly	Leu	Gly	Asn	Arg	Ser	Glu	
					150					155	;				160		
30	GCT	GTT	ATC	GCC	CAG	TTC	TAC	CAC	AGC	ССТ	GC0	GAC	AAC	AAG	CGG	TAC	611
	Ala	Val	He	Ala	Gla	Phe	Туг	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg	Tyr	
				165					170)				175			
35	ATC	TTT	GCA	GAC	GCT	TAT	GCC	CAG	TAC	стс	TGG	ATC	ACG	TTT	GAC	TTC	659
	lle	Phe	Ala	Asp	Ala	Tyr	Ala	Gln	Туг	Leu	Trp	He	Thr	Phe	Asp	Phe	
10			180					185					190				
••	TGC	AAC	ACT	CTT	CAA	GGC	TTT	TCC	ATC	CCA	TTT	CGG	GCA	GCT	GAT	CTC	707
	Cys	Asn	Thr	Leu	Gln	Gly	Phe	Ser	Ile	Pro	Phe	Arg	Ala	Ala	Asp	Leu	
15		195					200					205			~		
	CTC (CTA (CAC A	AGT A	AAG	GCC 1	rcc /	AAC	CTT	СТС	TTG	GGC 1	TTT (GAC A	AGG 1	CC	755
	Leu l	Leu F	lis S	Ser I	Lys i	Ala S	Ser A	Asn	Leu	Leu	Leu	Gly F	he A	Asp A	Arg S	Ser	
50	210					215					220					25	
•	CAC C	CCC A	AC A	AAG (CAG (CTG 1	rgg A	AAG	TCA	GAT	GAC	TTT C	GC C	CAG A	CC T	'GG	803

	His	Pro	Asn	Lys	Gln	Leu	Trp	Lys	Ser	Asp	Asp	Phe	Gly	Gln	Thr	Trp	
5					230					235					240		
	ATC	ATG	ATT	CAG	GAA	CAT	GTC	AAG	TCC	TTT	TCT	TGG	GGA	ATT	GAT	CCC	851
	He	Met	Пe	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Trp	Gly	lle	Asp	Pro	
10				245					250					255			
	TAT	GAC	AAA	CCA	AAT	ACC	ATC	TAC	ATT	GAA	CGA	CAC	GAA	CCC	тст	GGC	899
	Tyr	Asp	Lys	Pro	Asn	Thr	lle	Туг	Ile	Glu	Arg	His	Glu	Pro	Ser	Gly	
15			260					265					270				
	TAC	TCC	ACT	GTC	TTC	CGA	AGT	ACA	GAT	TTC	TTC	CAG	TCC	CGG	GAA	AAC	947
	Tyr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gln	Ser	Arg	Glu	Asn	
20		275					280					285					
	CAG	GAA	GTG	ATC	CTT	GAG	GAA	GTG	AGA	GAT	TTT	CAG	CTT	CGG	GAC	AAG	995
25	Gln	Glu	Val	lle	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp	Lys	
	290					295					300					305	
	TAC	ATG	TTT	GCT	ACA	AAG	GTG	GTG	CAT	CTC	TTG	GGC	AGT	GAA	CAG	CAG	1043
30	Tyr	Met	Phe	Ala	Thr	Lys	Vai	Val	His	Leu	Leu	Gly	Ser	Glu	Gln	Gln	
					310					315					320		
	TCT	тст	GTC	CAG	CTC	TGG	GTC	TCC	TTT	GGC	CGG	AAG	CCC	ATG	AGA	GCA	1091
35	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg	Ala	
				325					330					335			
40	GCC	CAG	TTT	GTC	ACA	AGA	CAT	CCT	ATT	AAT	GAA	TAT	TAC	ATC	GCA	GAT	1139
	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	He	Asn	Glu	Tyr	Tyr	Пе	Ala	Asp	
			340					345					350				
45	GCC	TCC	GAG	GAC	CAG	GTG	TTT	GTG	TGT	GTC	AGC	CAC	AGT	AAC	AAC	CGC	1187
	Ala	Ser	Glu	Asp	G1 n	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn	Arg	
		355					360					365					
50	ACC	AAT	TTA	TAC	ATC	TCA	GAG	GCA	GAG	GGG	СТG	AAG	TTC	TCC	CTG	TCC	1235
	Thr	Asn	Leu	Tyr	Пe	Ser	Glu	Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu	Ser	

	370				375					380					385	
_	TTG G	AG AAC	GTG	CTC	TAT	TAC	AGC	CCA	GGA	GGG	GCC	GGC	AGT	GAC	ACC	1283
5	Leu G	lu Asn	Val	Leu	Tyr	Tyr	Ser	Pro	Gly	Gly	Ala	Gly	Ser	Asp	Thr	
				390					395					400		
10	TTG G	TG AGG	TAT	TTT	GCA	AAT	GAA	CCA	TTT	GCT	GAC	TTC	CAC	CGA	GTG	1331
	Leu V	al Arg	Tyr	Phe	Ala	Asn	Glu	Pro	Phe	Ala	Asp	Phe	His	Arg	Val	
			405					410					415			
15	GAA G	GA TTG	CAA	GGA	GTC	TAC	ATT	GCT	ACT	CTG	ATT	AAT	GGT	TCT	ATG	1379
	Glu G	ly Leu	Gln	Gly	Val	Туг	Ile	Ala	Thr	Leu	Ile	Asn	Gly	Ser	Met	
20		420					425					430				
		AG GAG														1427
		lu Glu	Asn	Met	Arg		Val	He	Thr	Phe	Asp	Lys	Gly	Gly	Thr	
25		35	-	2.2	222	440					445					
		AG TTT														1475
		lu Phe	Leu	Gin		Pro	Ala	Phe	Thr		Tyr	Gly	Glu	Lys		
30	450		000	maa	455					460					465	
		GT GAG														1523
35	ASn C	ys Glu	Leu		GIN	GIY	Lys	Ser		HIS	Leu	Ala	GIN		Leu	
	ለርጥ ር	AG CTC	ርተር	470	ርጥር	CAC	ርጥር	CCC	475	A ጥር	ccc	ል ጥር	CTC	480	AAC	1571
		ln Leu														1571
40	50. 5	in bou	485		DCu	U 111	Dou	490	A1 8	MCC	110	116	495	561	Lys	
	GAG T	CG GCT		GGC	СТС	ATC	ATC		АСТ	GGC	тса	GTG		AAG	AAC	1619
45		er Ala														1010
		500					505			•		510		- •		
	TTG G	CT AGC	AAG	ACA	AAC	GTG	TAC	ATC	тст	AGC	AGT		GGA	GCC	AGG	1667
50	Leu A	la Ser	Lys	Thr	Asn	Val	Tyr	He	Ser	Ser	Ser	Ala	Gly	Ala	Arg	
		15				520					525		-			

	TGG	CGA	GAG	GCA	CTT	CCT	GGA	CCT	CAC	TAC	TAC	ACA	TGG	GGA	GAC	CAC	1715
5	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Tyr	Tyr	Thr	Trp	Gly	Asp	His	
	530					535					540					545	
	GGC	GGA	ATC	ATC	ACG	GCC	ATT	GCC	CAG	GGC	ATG	GAA	ACC	AAC	GAG	CTA ·	1763
10	Gly	Gly	lle	Ile	Thr	Ala	Ile	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu	Leu	
					550					555	•				560		
	AAA	TAC	AGT	ACC	AAT	GAA	GGG	GAG	ACC	TGG	AAA	ACA	TTC	ATC	TTC	TCT	1811
15	Lys	Tyr	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	Thr	Phe	Ile	Phe	Ser	
				565					570					575			
20	GAG	AAG	CCA	GTG	TTT	GTG	TAT	GGC	СТС	CTC	ACA	GAA	CCT	GGG	GAG	AAG	1859
	Glu	Lys	Pro	Val	Phe	Val	Tyr	Gly	Leu	Leu	Thr	Glu	Pro	Gly	Glu	Lys	
			580					585					590				
25	AGC	ACT	GTC	TTC	ACC	ATC	TTT	GGC	TCG	AAC	AAA	GAG	AAT	GTC	CAC	AGC	1907
	Ser	Thr	Val	Phe	Thr	Ιle	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His	Ser	
		595					600					605					
30	TGG	CTG	ATC	СТС	CAG	GTC	AAT	GCC	ACG	GAT	GCC	TTG	GGA	GTT	CCC	TGC	1955
	Τrp	Leu	lle	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Pro	Cys	
35	610					615					620					625	
55	ACA	GAG	AAT	GAC	TAC	AAG	CTG	TGG	TCA	CCA	тст	GAT	GAG	CGG	GGG	AAT	2003
	Thr	Glu	Asn	Asp	Tyr	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly	Asn	
40 .					630					635					640		
	GAG	TGT	TTG	CTG	GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	CGG	ACC	CCC	CAT	2051
	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro	His	
45				645					650					655			
	GCC	ACA	TGC	TTC	TAA	GGA	GAG	GAC	TTT	GAC	AGG	CCG	GTG	GTC	GTG	TCC	2099
	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val	Ser	
50			660					665					670				
	AAC	TGC	TCC	TGC	ACC	CGG	GAG	GAC	TAT	GAG	TGT	GAC	TTC	GGT	TTC	AAG	2147

	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Туг	Glu	Cys	Asp	Phe	Gly	Phe	Lys	
		675					680					685					
5	ATG	AGT	GAA	GAT	TTG	TCA	TTA	GAG	GTT	TGT	GTT	CCA	GAT	CCG	GAA	TTT	2195
	Met	Ser	Glu	Asp	Leu	Ser	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Glu	Phe	
10	690	,				695		,			700					705	•
	TCT	GGA	AAG	TCA	TAC	TCC	CCT	CCT	GTG	ССТ	TGC	CCT	GTG	GGT	тст	ACT	2243
	Ser	Gly	Lys	Ser	Туг	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Ser	Thr	
15					710					715					720		
	TAC	AGG	AGA	ACG	AGA	GGC	TAC	CGG	AAG	ATT	тст	GGG	GAC	ACT	TGT	AGC	2291
	Tyr	Arg	Arg	Thr	Arg	Gly	Tyr	Arg	Lys	Ιlе	Ser	Gly	Asp	Thr	Cys	Ser	
20				725					730					735			
	GGA	GGA	GAT	GTT	GAA	GCG	CGA	СТС	GAA	GGA	GAG	CTG	GTC	CCC	TGT	CCC	2339
25	Gly	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Cys	Pro	
25			740					745					750				
	CTG	GCA	GAA	GAG	AAC	GAG	TTC	ATT	CTG	TAT	GCT	GTG	AGG	AAA	TCC	ATC	2387
30	Leu	Ala	Glu	Glu	Asn	Glu	Phe	lle	Leu	Tyr	Ala	Val	Arg	Lys	Ser	lle	
		755	;				760					765	i				
	TAC	CGC	ТАТ	GAC	СТG	GCC	TCG	GGA	GCC	ACC	GAG	CAG	TTG	CCT	CTC	ACC	2435
3 5	Tyr	Arg	Tyr	Asp	Leu	Ala	Ser	Gly	Ala	Thr	Glu	Gln	Leu	Pro	Leu	Thr	
	770					775	i				780)				785	
	GGG	СТА	CGG	GCA	GCA	GTG	GCC	CTO	GAC	דדד	GAC	TAT	GAC	CAC	CAAC	TGT	2483
40	Gly	Leu	Arg	Ala	Ala	Val	Ala	Leu	Asp	Phe	. Asp	Туг	Glu	His	s Asr	Cys	
					790)				795	j .				800)	
45	TTG	ТАТ	r TGC	TCC	GAC	СТО	GCC	110	GAC	GTC	ATC	CAC	G CGC	СТО	TG1	TTG	2531
45	Leu	Туг	Trp	Ser	Asp	Leu	Ala	Leu	Asp	Val	I 1 e	e Glr	Arg	Lei	ı Cys	Leu	
				805	ō				810)				815	5 .		
50	AAT	r GGA	A AGC	ACA	GGC	G CAA	GAC	GTC	ATC	CATO	C AA1	r TC	r GGC	CTO	G GA	G ACA	2579
	Asr	ı Gly	y Ser	Thr	Gly	/ Glr	ı Glu	ı Val	He	e Ile	e Ası	n Sei	Gly	/ Ĺei	ı Glu	ı Thr	

			820					825					830)	•		
=	GTA	GAA	GCT	TTG	GCT	TTT	GAA	ccc	СТС	AGC	CAG	CTG	СТТ	TAC	TGG	GTA	2627
5	Val	Glu	Ala	Leu	Ala	Phe	Glu	Pro	Leu	Ser	Gln	Leu	Leu	Туг	Trp	Val	
		835					840					845					
10	GAT	GCA	GGC	TTC	AAA	AAG	ATT	GAG	GTA	GCT	AAT	CCA	GAT	GGC	GAC	TTC	2675
	Asp	Ala	Gly	Phe	Lys	Lys	Ile	Glu	Val	Ala	Asn	Pro	Asp	Gly	Asp	Phe	
	850					855					860					865	
15	CGA	CTC	ACA	ATC	GTC	AAT	TCC	TCT	GTG	CTT	GAT	CGT	CCC	AGG	GCT	CTG	2723
	Arg	Leu	Thr	Ile	Val	Asn	Ser	Ser	Val	Leu	Asp	Arg	Pro	Arg	Ala	Leu	
					870					875					880		
20	GTC	CTC	GTG	CCC	CAA	GAG	GGG	GTG	ATG	TTC	TGG	ACA	GAC	TGG	GGA	GAC	2771
	Val	Leu	Val	Pro	Gln	Glu	Gly	Val	Met	Phe	Тгр	Thr	Asp	Trp	Gly	Asp	
25				885					890					895			
	CTG	AAG	ССТ	GGG	ATT	TAT	CGG	AGC	ААТ	ATG	GAT	GGT	тст	GCT	GCC	TAT	2819
	Leu	Lys	Pro	Gly	Ile	Tyr	Arg	Ser	Asn	Met	Asp	Gly	Ser	Ala	Ala	Tyr	
30			900					905					910				
	CAC	CTG	GTG	TCT	GAG	GAT	GTG	AAG	TGG	CCC	AAT	GGC	ATC	TCT	GTG	GAC	2867
05	His	Leu	Val	Ser	Glu	Asp	Val	Lys	Тгр	Pro	Asn	Gly	lle	Ser	Val	Asp	
35		915					920		-			925					
	GAC	CAG	TGG	ATT	TAC	TGG	ACG	GAT	GCC	TAC	CTG	GAG	TGC	АТА	GAG	CGG	2915
40	Asp	Gln	Trp	Ile	Tyr	Trp	Thr	Asp	Ala	Tyr	Leu	Glu	Cys	Пе	Glu	Arg	
	930					935					940					945	
	ATC	ACG	TTC	AGT	GGC	CAG	CAG	CGC	тст	GTC	ATT	CTG	GAC	AAC	СТС	CCG	2963
45	Ile	Thr	Phe	Ser	Gly	G1 n	Gln	Arg	Ser	Val	Ιle	Leu	Asp	Asn	Leu	Pro	
					950					95 5					960		
	CAC	CCC	TAT	GCC	ATT	GCT	GTC	TTT	AAG	AAT	GAA	ATC	TAC	TGG	GAT	GAC	3011
50	His	Pro	Tyr	Ala	Пе	Ala	Val	Phe	Lys	Asn	Glu	Ile	Туг	Trp	Asp	Asp	
				965					970					975			

	TGG	TCA	CAG	CTC	AGC	ATA	TTC	CGA	GCT	TCC	AAA	TAC	AGT	GGG	TCC	CAG	3059
5	Trp	Ser	Gln	Leu	Ser	Ile	Phe	Arg	Ala	Ser	Lys	Туг	Ser	Gly	Ser	Gln	
•			980					985					990				
	ATG	GAG	АТТ	CTG	GCA	AAC	CAG	CTC	ACG	GGG	СТС	ATG	GAC	ATG	AAG	ATT	- 3107
10	Met	Glu	lle	Leu	Ala	Asn	Gln	Leu	Thr	Gly	Leu	Met	Asp	Met	Lys	lle	
		995					1000					1005	5				
	TTC	TAC	AAG	GGG	AAG	AAC	ACT	GGA	AGC	AAT	GCC	TGT	GTG	CCC	AGG	CCA	3155
15	Phe	Туг	Lys	Gly	Lys	Asn	Thr	Gly	Ser	Asn	Ala	Cys	Val	Pro	Arg	Pro	
	1010)				1015	i				1020)				1025	
	TGC	AGC	CTG	CTG	TGC	CTG	CCC	AAG	GCC	AAC	AAC	AGT	AGA	AGC	TGC	AGG	3203
20	Cys	Ser	Leu	Leu	Cys	Leu	Pro	Lys	Ala	Asn	Asn	Ser	Arg	Ser	Cys	Arg	
					1030)				103	5				104	0	
<i>25</i>	TGT	CCA	GAG	GAT	GTG	TCC	AGC	AGT	GTG	CTT	CCA	TCA	GGG	GAC	CTG	ATG	3251
25	Cys	Pro	Glu	Asp	Val	Ser	Ser	Ser	Val	Leu	Pro	Ser	Gly	Asp	Leu	Met	
		•		104					105					105			
30	TGT	GAC	TGC	CCT	CAG	GGC	ТАТ	CAG	СТС	AAG	AAC	AAT	ACC	TGT	GTC	AAA	3299
	Cys	Asp	Cys	Pro	Gln	Gly	Tyr	Gln	Leu	Lys	Asn	Asn	Thr	Cys	Val	Lys	
			106					106					107				
35	GAA	GAG	AAC	ACC	TGT	CTT	CGC	AAC	CAG	ТАТ	CGC	TGC	AGC	: AAC	GGC	AAC	3347
																Asn	
		107					108					108					
40	TGT			CAGO	: ATT	TGG	TGG	TGT	GAC	TT	GAC	CAAC	GAC	TG1	r GGA	GAC	3395
																/ Asp	
45	109					109					110					1105	5
45	ATC	a AGC	GA1	r GAC	G AGA	AAC	TGC	CC1	r ACO	CAC	CATO	C TG	r GA	CTC	G GA	CACC	3443
	Met	: Sei	r Ası	o Glu	ı Arg	g Asr	Cys	Pro	Th:	r Thi	110	е Су	s As	p Lei	i As	p Thr	
50					111						15				11		
	CAG	3 TT	r cg	T TG	C CAC	G GAC	TC1	r GG	G AC	T TG	T AT	c cc	A CT	G TC	C TA	T AAA	3491

	G	l n	Phe	Arg	у Су	s Gl	n Gl	u Se	r Gl	y Th	т Су	s II	e Pr	o Le	eu Se	er T	yr Lys	5
5					11	25				11	30				11	135		
																	AT TGT	
	C?	/S	Asp	Leu	Gli	ı As	p As	р Су	s Gl	y As	p As	n Se	r Ası	p GI	u Se	r H	is Cys	;
10				114					11-					11				
																	C ATG	
	Gl	u N	let	His	Glr	Cys	s Ar	g Se	r Ası	GI GI	и Ту	r Asr	Cys	Se	r Se	r Gl	y Met	
15			155					116					116					-
																	G GAC	
20	Cy.	s I	le.	Arg	Ser	Ser	Trp	o Val	Cys	: Ası	Gly	/ Asp	Asn	Asp	Cy:	s Ar	g Asp	
	11'						117					118					118	5
																	G GCC	3683
25	Tr	S	er <i>i</i>	Asp	Glu	Ala	Asn	Cys	Thr	Ala	Ile	Туг	His	Thr	Cys	s GI:	u Ala	
						119	0				119	5				120	00	
	TCC	A/	AC 1	TC	CAG	TGC	CGA	AAC	GGG	CAC	TGC	ATC	CCC	CAG	CGC	TG(G GCG	3731
30																	Ala	
					1205					121					121			
35	TGT	GA	C G	GG	GAT	ACG	GAC	TGC	CAG	GAT	GGT	TCC	GAT	GAG	GAT	CCA	GTC	3779
55																	Val	
				220					1225					1230				
10	AAC	TG	T G	AG A	AAG	AAG	TGC	AAT	GGA	TTC	CGC	TGC	CCA	AAC	GGC	ACT	TGC	3827
	Asn	Су	s G	lu (Lys	Lys	Cys	Asn	Gly	Phe	Arg	Cys	Pro	Asn	Gly	Thr	Cys	
		12	35					1240)				1245	,				
5	ATC	CC	A TI	CC A	GC .	AAA	CAT	TGT	GAT	GGT	CTG	CGT	GAT	TGC	TCT	GAT	GGC	3875
	Ile	Pr	o Se	er S	Ser :	Lys	His	Cys	Asp	Gly	Leu	Arg	Asp	Cys	Ser	Asp	Gly	
	1250						1255					1260					1265	
o	TCC	GA'	T GA	AA C	CAG (CAC	TGC	GAG	CCC	CTC	TGT	ACG (CAC	TTC	ATG	GAC	TTT	3923
	Ser	As	p GI	u G	iln i	His (Cys	Glu	Pro	Leu	Cys	Thr I	His I	Phe	Met	Asp	Phe	

					1270					1275	j				1280)	
5	GTG	TGT	AAG	AAC	CGC	CAG	CAG	TGC	CTG	TTC	CAC	TCC	ATG	GTC	TGT	GAC	3971
	Val	Cys	Lys	Asn	Arg	Gln	Gln	Cys	Leu	Phe	His	Ser	Met	Val	Cys	Asp	
				1285	,				1290)				1299	5		
10	GGA	ATC	ATC	CAG	TGC	CGC	GAC	GGG	TCC	GAT	GAG	GAT	GCG	GCG	TTT	GCA	4019
	Gly	Ile	lle	Gln	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Asp	Ala	Ala	Phe	Ala	
			1300)				1305	5				1310)			
15	GGA	TGC	TCC	CAA	GAT	CCT	GAG	TTC	CAC	AAG	GTA	TGT	GAT	GAG	TTC	GGT	4067
	Gly	Cys	Ser	Gln	Asp	Pro	Glu	Phe	His	Lys	Val	Cys	Asp	Glu	Phe	Gly	
20		131	5				1320)				132	5				
20																GAC	4115
	Phe	Gln	Cys	Gln	Asn	Gly	Val	Cys	He	Ser	Leu	Ile	Trp	Lys	Cys	Asp	
25	133					133					134					1345	
																AAC	4163
	Gly	Met	Asp	Asp	Cys	Gly	Asp	Туг	Ser	Asp	Glu	Ala	Asn	Cys		ı Asn	
30					135					135						50	4011
																r GAG	4211
05	Pro	Thr	Glu	Ala	Pro	Asn	Cys	Ser	Arg	Туг	Phe	e Glr	Phe			s Glu	
35				136					137					13'			4050
																C GAC	4259
40	Asr	Gly	His	Cys	116	Pro	Asn	Arg	Tr) Ly:	s Cys	s Ası			u As:	n Asp	
			138					138					139			m 0000	4207
																T CTT	4307
45	Cys	s Gl	y Ası	o Trg	Sei	r Asp			s As	р Су	s Gl			r Hi	S II	e Leu	
		139					140					14			C MA	C CCC	4355
																.C CGC	4555
50			e Se	r Th	r Pr) Sei	r Th	гСу			U AS	пту	т ту	r Arg	;
	14	10				14	15				14	20				1425	į.

	TGO	CAG	C AG	T GG	G AC	C TG	C GT	G AT	G GA	C AC	C TO	GG G	TG T	GC (GAC	GGG	TAC	4403
5	Cys	s Se	r Se	r Gl	y Th	r Cy:	s Va	l Me	t As	p Th	ır Tr	p V.	al C	ys i	Asp	Glv	Туг	4400
Ü					14						35				•	144		_
	CGA	GA'	T TG	T GC	A GA	r GG(C TC1	Γ GA	C GA	G GA	A GC	C TO	GC C	CC 1	rtg	CTT	GCA	4451
10	Arg	, Ası	p Cy	s Al	a Ası	o Gly	/ Ser	As	p Gl	u Gl	u Al	a Cy	/s P	ro L	Leu	Leu	Ala	4401
				14					14						455			
	AAC	GTO	CAC	r GC	r GCO	TCC	ACT	, cc	CAC	C CA	A CT	T GO	G CO				CGA	4499
15	Asn	Val	Thr	Ala	a Ala	Ser	Thr	Pro	Th:	r Gli	n Le	u Gl	y Aı	g C	ys	Asp	Arg	4400
			146					146						170	•		6	
	TTT	GAG	TTC	GAA	TGC	CAC	CAA	CCC	AA(ACC	G TG	г ат			AC	TGG	AAG	4547
20	Phe	Glu	Phe	Glu	Cys	His	Gln	Pro	Lys	Thr	· Cys	s II	e Pr	o A	sn	Тгр	Lvs	1041
		147					148				•	14				•	, -	
25	CGC	TGT	GAC	GGC	CAC	CAA	GAT	TGC	CAG	GAT	GGC	C CG	G GA	C G	AG	GCC	ААТ	4595
	Arg	Cys	Asp	Gly	His	Gln	Asp	Cys	Gln	Asp	Gly	/ Аг	g As	p G	lu .	Ala	Asn	1000
	1490					1495					150						1505	
30	TGC	CCC	ACA	CAC	AGC	ACC	TTG	ACT	TGC	ATG	AGC	: AG(G GA	G TY	rc (CAG		4643
	Cys																	.010
					1510					151						520	. , .	
35	GAG	GAC	GGG	GAG	GCC	TGC	АТТ	GTG	CTC			CGC	TG(C GA			LTC	4691
	Glu																	4001
40				1525					1530				•	15			•••	
	CTG (GAC	TGC	TCG	GAC	GAG	AGC	GAT	GAA	AAG	GCC	TGC	AG1			ag 1	TG	4739
	Leu i																	
45			1540					1545				·	155					
	ACT (GTG	TAC	AAA	GTA	CAG I	AAT (CTT	CAG	TGG	ACA	GCT			СТ	CT G	:CC	4787
	Thr \																	1,01
50		1555					1560			_		156			_		- ,	
	GAT (GTG .	ACT	TTG	ACC '	rgg <i>i</i>	ATG A	AGG -	CCC	AAA	AAA			TC1	r G	CA T	CT	4835
												_					- •	1000

	Asp	Val	Thr	Leu	Thr	Trp	Met	Arg	Pro	Lys	Lys	Met	Pro	Ser	Ala	Ser	
	1570)				1575	5				1580	1				1585	
5	TGT	GTA	TAT	AAT	GTC	TAC	TAC	AGG	GTG	GTT	GGA	GAG	AGC	АТА	TGG	AAG	4883
	Cys	Val	Туг	Asn	Val	Туг	Tyr	Arg	Val	Val	Gly	Glu	Ser	He	Trp	Lys	
10					1590)				1595	5				1600)	
	ACT	CTG	GAG	ACC	CAC	AGC	AAT	AAG	ACA	AAC	ACT	GTA	TTA	AAA	GTC	TTG	4931
	Thr	Leu	Glu	Thr	His	Ser	Asn	Lys	Thr	Asn	Thr	Val	Leu	Lys	Val	Leu	
15				1605	ō				161	C				161	5		
	AAA	CCA	GAT	ACC	ACG	ТАТ	CAG	GTT	AAA	GTA	CAG	GTT	CAG	TGT	CTC	AGC	4979
	Lys	Pro	Asp	Thr	Thr	Tyr	Gln	Val	Lys	Val	Gln	Val	Gln	Cys	Leu	Ser	
20			1620	0				1629	5				1630	0			
	AAG	GCA	CAC	AAC	ACC	ААТ	GAC	TTT	GTG	ACC	CTG	AGG	ACC	CCA	GAG	GGA	5027
25	Lys	Ala	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly	
		163	5				164	0				164	5				
	TTG	CCA	GAT	GCC	CCT	CGA	AAT	CTC	CAG	CTG	TCA	CTC	CCC	AGG	GAA	GCA	5075
30	Leu	Pro	Asp	Ala	Pro	Arg	Asn	Leu	Gln	Leu	Ser	Leu	Pro	Arg	Glu	Ala	
	1650	0				165	5				166	0				1665	
	GAA	GGT	GTG	ATT	GTA	GGC	CAC	TGG	GCT	CCT	CCC	ATC	CAC	ACC	CAT	GGC	5123
35	Glu	Gly	Val	Ile	Val	Gly	His	Trp	Ala	Pro	Pro	He	His	Thr	His	Gly	
					167	0				167	5				168	0	
	CTC	ATC	CGT	GAG	TAC	ATT	GTA	GAA	TAC	AGC	AGG	AGT	GGT	TCC	AAG	ATG	5171
40	Leu	Ile	Arg	Glu	Туг	He	Val	Glu	Туr	Ser	Arg	Ser	Gly	Ser	Lys	Met	
				168	5				169	0				169	5		
45	TGG	GCC	TCC	CAG	AGG	GCT	GCT	AGT	AAC	TTT	ACA	GAA	ATC	AAG	AAC	TTA	5219
	Trp	Ala	Ser	Gln	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	lle	Lys	Asn	Leu	
			170	0				170	5				171	0			
50	TTG	GTC	AAC	ACT	СТА	TAC	ACC	GTC	AGA	GTG	GCT	GCG	GTG	ACT	AGT	CGT	5267
	Leu	Val	Asn	Thr	Leu	Туг	Thr	Val	Arg	Val	Ala	Ala	Val	Thr	Ser	Arg	

		171	.5				172	0				172	5				
5	GGA	. ATA	GGA	AAC	TGG	AGC	GAT	ТСТ	AAA	TCC	ATT	ACC	ACC	АТА	AAA	GGA	5315
•													Thr				
	173					173					174					1745	
10	AAA	GTG	ATC	CCA	CCA	CCA	GAT	ATC	CAC	ATT	GAC	AGC	ТАТ	GGT	GAA		5363
													Tyr				
					175					175					176		
15	TAT	СТА	AGC	TTC	ACC	CTG	ACC	ATG	GAG	AGT	GAT	ATC	AAG	GTG	AAT	GGC	5411
	Tyr	Leu	Ser	Phe	Thr	Leu	Thr	Met	Glu	Ser	Asp	Ile	Lys	Val	Asn	Gly	
				176	5				177	0				177	5		
20	TAT	GTG	GTG	AAC	CTT	TTC	TGG	GCA	TTT	GAC	ACC	CAC	AAG	CAA	GAG	AGG	5459
	Tyr	Val	Val	Asn	Leu	Phe	Trp	Ala	Phe	Asp	Thr	His	Lys	Gln	Glu	Arg	
25			1780)				178	5				1790)			
	AGA	ACT	TTG	AAC	TTC	CGA	GGA	AGC	ATA	TTG	TCA	CAC	AAA	GTT	GGC	AAT	5507
	Arg	Thr	Leu	Asn	Phe	Arg	Gly	Ser	He	Leu	Ser	His	Lys	Val	Gly	Asn	
30		179	5				1800)				1805	5				
	CTG	ACA	GCT	CAT	ACA	TCC	ТАТ	GAG	ATT	TCT	GCC	TGG	GCC	AAG	ACT	GAC	5555
	Leu	Thr	Ala	His	Thr	Ser	Tyr	Glu	lle	Ser	Ala	Trp	Ala	Lys	Thr	Asp	
35	1810)				1815	5				1820)				1825	
	TTG	GGG	GAT	AGC	CCT	CTG	GCA	TTT	GAG	CAT	GTT	ATG	ACC	AGA	GGG	GTT	5603
40	Leu	Gly	Asp	Ser	Pro	Leu	Ala	Phe	Glu	His	Val	Met	Thr	Arg	Gly	Val	
40					1830)				1835	5				1840)	
	CGC	CCA	CCT	GCA	CCT	AGC	СТС	AAG	GCC	AAA	GCC	ATC	AAC	CAG	ACT	GCA	5651
45	Arg	Pro	Pro	Ala	Pro	Ser	Leu	Lys	Ala	Lys	Ala	Пе	Asn	Gln	Thr	Ala .	
				1845	i				1850)				1855	i		
	GTG	GAA	TGT	ACC	TGG	ACC	GGC	CCC	CGG	ААТ	GTG	GTT	TAT	GGT	ATT	TTC	5699
50	Val	Glu	Cys	Thr	Trp	Thr	Gly	Pro	Arg	Asn	Val	Val	Tyr	Gly	Пе	Phe	
			1860)				1865					1870	1			

	TAT	GCC	ACG	TCC	TTT	CTT	GAC	СТС	TAT	CGC	AAC	CCG	AAG	AGC	ŤТG	ACT	•	5747
	Tyr	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Tyr	Arg	Asn	Pro	Lys	Ser	Leu	Thr		
5		1875	5				1880)				1885	ó					
	ACT	TCA	CTC	CAC	AAC	AAG	ACG	GTC	ATT	GTC	AGT	AAG	GAT	GAG	CAG	TAT		5795
10	Thr	Ser	Leu	His	Asn	Lys	Thr	Val	Ile	Val	Ser	Lys	Asp	Glu	Gln	Tyr		
	1890)				1895	5				1900)				1905		
	TTG	TTT	CTG	GTC	CGT	GTA	GTG	GTA	CCC	TAC	CAG	GGG	CCA	TCC	TCT	GAC		5843
15	Leu	Phe	Leu	Val	Arg	Val	Val	Val	Pro	Туr	Gln	Gly	Pro	Ѕег	Ser	Asp		
					1910)				1915	5				1920)		
	TAC	CTT	GTA	GTG	AAG	ATG	ATC	CCG	GAC	AGC	AGG	CTT	CCA	CCC	CGT	CAC		5891
20	Tyr	Val	Val	Val	Lys	Met	Ile	Pro	Asp	Ser	Arg	Leu	Pro	Pro	Arg	His		
				1929	ō				1930)				1935	5			
25	CTG	CAT	GTG	GTT	CAT	ACG	GGC	AAA	ACC	TCC	GTG	GTC	ATC	AAG	TGG	GAA		5939
	Leu	His	Val	Val	His	Thr	Gly	Lys	Thr	Ser	Val	Val	Ile	Lys	Trp	Glu		
			194	0				194	5				195	0				
30	TCA	CCG	TAT	GAC	TCT	CCT	GAC	CAG	GAC	TTG	TTG	TAT	GCA	ATT	GCA	GTC		5987
	Ser	Рго	Tyr	Asp	Ser	Pro	Asp	Gln	Asp	Leu	Leu	Tyr	Ala	He	Ala	Val		
		195	5				196	0				196	5					
35	AAA	GAT	CTC	ATA	AGA	AAG	ACT	GAC	AGG	AGC	TAC	AAA	GTA	AAA	TCC	CGT		6035
	Lys	Asp	Leu	He	Arg	Lys	Thr	Asp	Arg	Ser	Туг	Lys	Val	Lys	Ser	Arg		
40	197	0				197	5				198	0				1985	j	
40							100	CAAL	AAC	ΔΔΩ	ተጥር	CAC	CCT	CCC	CCC	ΔΔΔ		6083
	AAC	AGC	ACT	GTG	GAA	TAC	ACC	CII	AAC	nnu	110	unu	CUI	UUL	ննն	nnn		
								Leu										
45						Туг					Leu					Lys		
45	Asn	Ser	Thr	Val	Gl u	Туг 10	Thr	Leu	Asn	Lys	Leu 15	Glu	Pro	Gly	Gly 200	Lys		6131
45	Asn TAC	Ser	Thr	Val	Glu 199 GTC	Туг 90 САА	Thr	Leu G GGG	Asn AAC	Lys 199 ATC	Leu 5 AGC	Glu C AAA	Pro	Gly	Gly 200 AGC	Lys 0		6131
45	Asn TAC	Ser	Thr	Val	Glu 199 GTC Val	Туг 90 САА	Thr	Leu G GGG	Asn AAC	Lys 199 ATG	Leu 5 AGC	Glu C AAA	Pro	Gly	Gly 200 AGC Ser	Lys O ATA		6131

	Lys	Ile	Thr	Thr	Val	Ser	Leu	Ser	Ala	a Pro	Asp	Ala	Leu	Lys	lle	Ile	
5			202	0				202	25				203	0			
	ACA	GAA	AAT	GAT	CAT	GTT	CTT	CTO	777	r TGC	AAA	AGC	CTG	GCT	TTA	AAG	6227
	Thr	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys	
10		203	5				204	0				204	5		-		
	GAA	AAG	CAT	TTT	AAT	GAA	AGC	AGG	GGC	ТАТ	GAG	ATA	CAC	ATG	TTT	GAT	6275
	Glu	Lys	His	Phe	Asn	Glu	Ser	Arg	Gly	T yr	Glu	He	His	Met	Phe	Asp	
15	205	0				205	5				206	0				2065	
	AGT	GCC	ATG	AAT	ATC	ACA	GCT	TAC	CTT	GGG	ААТ	ACT	АСТ	GAC	ААТ	TTC	6323
20	Ser	Ala	Met	Asn	Ile	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn	Phe	
					2070)				207	5				2080)	
	TTT	AAA	ATT	TCC	AAC	CTG	AAG	ATG	GGT	CAT	AAT	TAC	ACG	TTC	ACC	GTC	6371
25	Phe	Lys	Пе	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Phe	Thr	Val	
				2085	j				209	0				209	5		
	CAA	GCA	AGA	TGC	CTT	TTT	GGC	AAC	CAG	ATC	TGT	GGG	GAG	ССТ	GCC	ATC	6419
30	Gln	Ala	Arg	Cys	Leu	Phe	Gly	Asn	Gln	lle	Cys	Gly	Glu	Pro	Ala	He	
			2100)				2105	5				2110)			
35	CTG	CTG	TAC	GAT	GAG	CTG	GGG	тст	GGT	GCA	GAT	GCA	тст	GCA	ACG	CAG	6467
	Leu	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Glŷ	Ala	Asp	Ala	Ser	Ala	Thr	G1 n	
		2115					2120	١				2125	5			•	
40	GCT	GCC	AGA	TCT	ACG	GAT	GTT	GCT	GCT	GTG	GTG	GTG	CCC	ATC	TTA	TTC	6515
	Ala	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	lle	Leu	Phe	
	2130	i				2135	j				2140)				2145	
45	CTG	ATA	CTG	CTG	AGC	CTG	GGG	GTG	GGG	TTT	GCC	ATC	CTG	TAC	ACG	AAG	6563
	Leu	lle	Leu	Leu	Ser	Leu	Gly	Val	Gly	Phe	Ala	He	Leu	Tyr	Thr	Lys	
					2150					2155	5				2160		
50	CAC	CGG	AGG	CTG	CAG	AGC	AGC	TTC	ACC	GCC	TTC	GCC	AAC	AGC	CAC	TAC	6611
	His	Arg	Arg	Leu	Gln	Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	His	Tyr	

				2165	5				2170)				2179	5		
5	AGC	TCC	AGG	CTG	GGG	TCC	GCA	ATC	TTC	TCC	TCT	GGG	GAT	GAC	CTG	GGG	6659
	Ser	Ser	Arg	Leu	Gly	Ser	Ala	Ile	Phe	Ser	Ser	Gly	Asp	Asp	Leu	Gly	
			2180)				218	5				219	0			
10	GAA	GAT	GAT	GAA	GAT	GCC	CCT	ATG	ATA	ACT	GGA	TTT	TCA	GAT	GAC	GTC	6707
	Glu	Asp	Asp	Glu	Asp	Ala	Pro	Met	Ile	Thr	Gly	Phe	Ser	Asp	Asp	Val	
15		219	5				2200)				220	5				
	CCC	ATG	GTG	ATA	GCC	TGA	AAGA	GCT 1	TTCC'	rcac	TA G	AAAC	CAAA	T GG	TGTA	AATA	6762
	Pro	Met	Val	He	Ala												
20	2210)															
	TTT	ratt	rga 1	γααλ(GATAC	GT TO	GATG	GTTT/	1 TT	TAA	AAGA	TGC	ACTT	TGA (GTTG	CAATAT	6822
25	GTTA	ATTT	TTA 1	ratg(GGCC/	AA A											6843
00																	
<i>30</i>																	
35																	
40																	
45																	
50																	
55																	

SEQUENCE LISTING

5	(1) GENERAL INFORMATION:	
10	 (i) APPLICANT: (A) NAME: KOWA CO., LTD. (B) STREET: 6-29, Nishiki 3-chome, Naka-ku, Nagoya-shi, (C) CITY: Aichi (E) COUNTRY: Japan (F) POSTAL CODE (ZIP): none 	
	(ii) TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND TH GENE CODING THEREFOR	Œ
15	(iii) NUMBER OF SEQUENCES: 7 (iv) COMPUTER READABLE FORM:	
22	(A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)	
20	(2) INFORMATION FOR SEQ ID NO: 1:	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6639 base pairs (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: cDNA to mRNA	
30		
•	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
	ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCCT TCCTATTCAC CCTGGTCGCG	60
35	CTGCTGCCGC CCGGGGCTCT CTGCGAGGTG TGGACGCGGA CACTGCACGG CGGCCGCGCG	120
	CCCTTACCCC AGGAGCGGGG CTTCCGCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG	180
•	TGGGAGCGCG GGGATGCCAG GGGGGGAGC CGGGGGACG AGAAGCCGCT CCGGAGGAGA	240
40	CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTCAG CCTCAATGAT	300
	TCCCACAATC AGATGGTGGT GCACTGGCCC GGAGAGAAAA GCAACGTGAT CGTGGCCTTG	360
	GCCCGGGACA GCCTGGCGTT GGCCAGGCCC AGGAGCAGTG ATGTGTACGT GTCTTATGAC	420
45	TATGGAAAAT CATTCAATAA GATTTCAGAG AAATTGAACT TCGGCGCGGG AAATAACACA	480
	GAGGCTGTGG TGGCCCAGTT CTACCACAGC CCTGCGGACA ACAAACGGTA CATCTTCGCA	540
	GATGCCTACG CCCAGTATCT CTGGATCACG TTTGACTTCT GCAACACCAT CCATGGCTTT	600
50	TCCATCCCGT TCCGGGCAGC TGATCTCCTA CTCCACAGTA AGGCCTCCAA CCTTCTCCTG	660
	GGCTTCGACA GGTCTCACCC CAACAAGCAG CTGTGGAAGT CGGATGATTT TGGCCAGACC	720
	TGGATCATGA TTCAAGAACA CGTGAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA	780

68

	CCAAACACCA	TCTACATCGA	ACGGCACGAA	CCTTCTGGCT	ACTCCACGGT	TTTCCGAAGT	840
	ACAGACTTCT	TCCAGTCCCG	GGAAAA.CCAG	GAAGTGATCT	TGGAGGAAGT	GAGAGACTTT	900
5	CAGCTTCGGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTCCACTG	960
	CAGTCTTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCGGAAGC	CCATGCGGGC	CGCCCAGTTT	1020
	GTTACAAGAC	ATCCTATCAA	CGAATATTAC	ATCGCGGATG	CCTCGGAGGA	CCAGGTGTTT	1080
10	GTGTGTGTCA	GTCACAGCAA	CAACCGCACC	AACCTCTACA	TCTCGGAGGC	AGAGGGCTTG	1140
	AAGTTCTCTC	TGTCCCTGGA	GAACGTGCTC	TACTACACCC	CGGGAGGGC	CGGCAGTGAC	1200
	ACCTTGGTGA	GGTACTTTGC	AAATGAACCG	TTTGCTGACT	TCCATCGTGT	GGAAGGGTTG	1260
15	CAGGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCT	1320
	GTCATCACCT	TTGACAAAGG	GGGCACCTGG	GAATTTCTGC	AGGCTCCAGC	CTTCACGGGG	1380
	TATGGAGAGA	AAATCAACTG	TGAGCTGTCC	GAGGGCTGTT	CCCTCCACCT	GGCCCAGCGC	1440
20	CTCAGCCAGC	TGCTCAACCT	CCAGCTCCGG	AGGATGCCCA	TCCTGTCCAA	GGAGTCGGCG	1500
	CCTGGCCTCA	TCATTGCCAC	GGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	TCACTACTAT	1620
25	ACATGGGGAG	ACCATGGCGG	CATCATCATG	GCCATTGCCC	AAGGCATGGA	AACCAACGAA	1680
	CTGAAGTACA	GTACCAACGA	AGGGGAGACC	TGGAAAGCCT	TCACCTTCTC	TGAGAAGCCC	1740
	GTGTTTGTGT	ATGGGCTCCT	CACGGAACCC	GGCGAGAAGA	GCACGGTCTT	CACCATCTTT	1800
30	GGCTCCAACA	AGGAGAACGT	GCACAGCTGG	CTCATCCTCC	AGGTCAATGC	CACAGACGCC	1860
	CTGGGGGTTC	CTTGCACAGA	GAACGACTAC	AAGCTCTGGT	CACCATCTGA	TGAGCGGGGG	1920
	AATGAGTGTT	TGCTTGGACA	CAAGACTGTT	TTCAAACGGA	GGACCCCGCA	CGCCACATGC	1980
35	TTTAACGGAG	AAGACTTTGA	CAGGCCGGTG	GTTGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
	GACTATGAGT	GTGACTTTGG	CTTCCGGATG	AGTGAAGACT	TGGCATTAGA	GGTGTGTGTT	2100
	CCAGATCCAG	GATTTTCTGG	AAAGTCCTCC	CCTCCAGTGC	CTTGTCCCGT	GGGCTCTACG	2160
	TACAGGCGAT	CAAGAGGCTA	CCGGAAGATT	TCTGGGGACA	CCTGTAGTGG	AGGAGATGTT	2220
40	GAGGCACGGC	TAGAAGGAGA	GCTGGTCCCC	TGTCCCCTGG	CAGAAGAGAA	CGAGTTCATC	2280
	CTGTACGCCA	CGCGCAAGTC	CATCCACCGC	TATGACCTGG	CTTCCGGAAC	CACGGAGCAG	2340
	TTGCCCCTCA	CTGGGTTGCG	GGCAGCAGTG	GCCCTGGACT	TTGACTATGA	GCACAACTGC	2400
4 5	CTGTATTGGT	CTGACCTGGC	CTTGGACGTC	ATCCAGCGCC	TCTGTTTGAA	CGGGAGTACA	2460
	GGACAAGAGG	TGATCATCAA	CTCTGACCTG	GAGACGGTAG	AAGCTTTGGC	TTTTGAACCC	2520
	CTCAGCCAAT	TACTTTACTG	GGTGGACGCA	GGCTTTAAAA	AGATCGAGGT	AGCCAATCCA	2580
50	GATGGTGACT	TCCGACTCAC	CGTCGTCAAT	TCCTCGGTGC	TGGATCGGCC	CCGGGCCCTG	2640
	CTCCTTCTC	CCCDDGDDGG	САТСАТСТТС	тссасссаст	GGGGAGACCT	GAAGCCTGGG	2700

	ATTTATCGGA	GCAACATGGA	CGGATCTGCC	GCCTATCGCC	TCGTG CCGGA	GG.ATGTGAAG	2760
5	TGGCCCAATG	GCATTTCCGT	GGACGATCAG	TGGATCTACT	GGACGGATGC	CTACCTGGAC	2820
	TGCATTGAGC	GCATCACGTT	CAGCGGCCAG	CAGCGCTCCG	TCATCCTGGA	CAGACTCCCG	2880
	CACCCCTATG	CCATTGCTGT	CTTTAAGAAT	GAGATTTACT	GGGATGACTG	GTCACAGCTC	2940
	AGCATATTCC	GAGCTTCTAA	GTACAGCGGG	TCCCAGATGG	AGATTCTGGC	CAGCCAGCTC	3000
10	ACGGGGCTGA	TGGACATGAA	GATCTTCTAC	AAGGGGAAGA	ACACAGGAAG	CAATGCGTGT	3060
	GTACCCAGGC	CGTGCAGCCT	GCTGTGCCTG	CCCAGAGCCA	ACAACAGCAA	AAGCTGCAGG	3120
	TGTCCAGATG	GCGTGGCCAG	CAGTGTCCTC	CCTTCCGGGG	ACCTGATGTG	TGACTGCCCT	3180
15	AAGGGCTACG	AGCTGAAGAA	CAACACGTGT	GTCAAAGAAG	AAGACACCTG	TCTGCGCAAC	3240
	CAGTACCGCT	GCAGCAACGG	GAACTGCATC	AACAGCATCT	GGTGGTGCGA	TTTCGACAAC	3300
	GACTGCGGAG	ACATGAGCGA	CGAGAAGAAC	TGCCCTACCA	CCATCTGCGA	CCTGGACACC	3360
20	CAGTTCCGTT	GCCAGGAGTC	TGGGACGTGC	ATCCCGCTCT	CCTACAAATG	TGACCTCGAG	3420
	GATGACTGTG	GGGACAACAG	TGACGAAAGG	CACTGTGAAA	TGCACCAGTG	CCGGAGCGAC	3480
	GAATACAACT	GCAGCTCGGG	CATGTGCATC	CGCTCCTCCT	GGGTGTGCGA	CGGGGACAAC	3540
25	GACTGCAGGG	ACTGGTCCGA	CGAGGCCAAC	TGCACAGCCA	TCTATCACAC	CTGTGAGGCC	3600
	TCCAACTTCC	AGTGCCGCAA	CGGGCACTGC	ATCCCCCAGC	GGTGGGCGTG	TGACGGCGAC	3660
	GCCGACTGCC	AGGATGGCTC	TGATGAGGAT	CCAGCCAACT	GTGAGAAGAA	GTGCAACGGC	3720
30	TTCCGCTGCC	CGAACGGCAC	CTGCATTCCC	TCCACCAAGC	ACTGTGACGG	CCTGCACGAT	3780
	TGCTCGGACG	GCTCCGACGA	GCAGCACTGC	GAGCCCCTGT	GTACACGGTT	CATGGACTTC	3840
	GTGTGTAAGA	ACCGCCAGCA	GTGCCTCTTC	CACTCCATGG	TGTGCGATGG	GATCATCCAG	3900
35	TGCCGTGACG	GCTCCGACGA	GGACCCAGCC	TTTGCAGGAT	GCTCCCGAGA	CCCCGAGTTC	3960
	CACAAGGTGT	GCGATGAGTT	CGGCTTCCAG	TGTCAGAACG	GCGTGTGCAT	CAGCTTGATC	4020
	TGGAAGTGCG	ACGGGATGGA	TGACTGCGGG	GACTACTCCG	ACGAGGCCAA	CTGTGAAAAC	4080
10	CCCACAGAAG	CCCCCAACTG	CTCCCGCTAC	TTCCAGTTCC	GGTGTGACAA	TGGCCACTGC	4140
	ATCCCCAACA	GGTGGAAGTG	TGACAGGGAG	AATGACTGTG	GGGACTGGTC	CGACGAGAAG	4200
	GACTGTGGAG	ATTCACATGT	ACTTCCGTCT	ACGACTCCTG	CACCCTCCAC	GTGTCTGCCC	4260
4E	AATTACTACC	GCTGCGGCGG	GGGGGCCTGC	GTGATAGACA	CGTGGGTTTG	TGACGGGTAC	4320
1 5	CGAGATTGCG	CAGATGGATC	CGACGAGGAA	GCCTGCCCCT	CGCTCCCCAA	TGTCACTGCC	4380
	ACCTCCTCCC	CCTCCCAGCC	TGGACGATGC	GACCGATTTG	AGTTTGAGTG	CCACCAGCCA	4440
	AAGAAGTGCA	TCCCTAACTG	GAGACGCTGT	GACGGCCATC	AGGATTGCCA	GGATGGCCAG	4500
50	GACGAGGCCA	ACTGCCCCAC	TCACAGCACC	TTGACCTGCA	TGAGCTGGGA	GTTCAAGTGT	4560
	GAGGATGGCG	AGGCCTGCAT	CGTGCTGTCA	GAACGCTGCG	ACGGCTTCCT	GGACTGCTCA	4620

	GATGAGAGCG	ACGAGAAGGC	CTGCAGTGAT	GAGTTAACTG	TA FACA AAGT	ACAGAATCTT	4680
	CAGTGGACAG	CTGACTTCTC	TGGGAATGTC	ACTTTGACCT	GGATGCGGCC	CAAAAAATG	4740
5	CCCTCTGCTG	CTTGTGTATA	CAACGTGTAC	TATAG.4GTTG	TTGGAGAGAG	CATATGGAAG	4800
	ACTCTGGAGA	CTCACAGCAA	TAAGACAAAC	ACTGTATTAA	AAGTGTTGAA	ACCAGATACC	4860
	ACCTACCAGG	TTAAAGTGCA	GGTTCAGTGC	CTGAGCAAGG	TGCACAACAC	CAATGACTTT	4920
10	GTGACCTTGA	GAACTCCAGA	GGGATTGCCA	GACGCCCCTC	AGAACCTCCA	GCTGTCGCTC	4980
	CACGGGGAAG	AGGAAGGTGT	GATTGTGGGC	CACTGGAGCC	CTCCCACCCA	CACCCACGGC	5040
	CTCATTCGCG	AATACATTGT	AGAGTATAGC	AGGAGTGGTT	CCAAGGTGTG	GACTTCAGAA	5100
15	AGGGCTGCTA	GTAACTTTAC	AGAAATAAAG	AACTTGTTGG	TCAACACCCT	GTACACCGTC	5160
	AGAGTGGCTG	CGGTGACGAG	TCGTGGGATA	GGAAACTGGA	GCGATTCCAA	ATCCATTACC	5220
	ACCGTGAAAG	GAAAAGCGAT	CCCGCCACCA	AATATCCACA	TTGACAACTA	CGATGAAAAT	5280
20	TCCCTGAGTT	TTACCCTGAC	CGTGGATGGG	AACATCAAGG	TGAATGGCTA	TGTGGTGAAC	5340
	CTTTTCTGGG	CATTTGACAC	CCACAAACAA	GAGAAGAAAA	CCATGAACTT	CCAAGGGAGC	5400
	TCAGTGTCCC	ACAAAGTTGG	CAATCTGACA	GCACAGACGG	CCTATGAGAT	TTCCGCCTGG	5460
25	GCCAAGACTG	ACTTGGGCGA	TAGTCCTCTG	TCATTTGAGC	ATGTCACGAÇ	CAGAGGGGTT	5520
23	CGCCCACCTG	CTCCTAGCCT	CAAGGCCAGG	GCTATCAATC	AGACTGCAGT	GGAATGCACC	5580
	TGGACAGGCC	CCAGGAATGT	GGTGTATGGC	ATTTTCTATG	CCACATCCTT	CCTGGACCTC	5640
	TACCGCAACC	CAAGCAGCCT	GACCACGCCG	CTGCACAACG	CAACCGTGCT	CGTCGGTAAG	5700
30	GATGAGCAGT	ATCTGTTTCT	GGTCCGGGTG	GTGATGCCCT	ACCAAGGGCC	GTCCTCGGAC	5760
	TACGTGGTCG	TGAAGATGAT	CCCGGACAGC	AGGCTTCCTC	CCCGGCACCT	GCATGCCGTT	5820
	CACACCGGCA	AGACCTCGGC	CGTCATCAAG	TGGGAGTCGC	CCTACGACTC	TCCTGACCAG	5880
35	GACCTGTTCT	ATGCGATCGC	AGTTAAAGAT	CTGATACGAA	AGACGGACCG	GAGCTACAAA	5940
	GTCAAGTCCC	GCAACAGCAC	CGTGGAGTAC	ACCCTGAGCA	AGCTGGAGCC	CGGAGGGAAA	6000
	TACCACGTCA	TTGTGCAGCT	GGGGAACATG	AGCAAAGATG	CCAGTGTGAA	GATCACCACC	6060
40	GTTTCGTTAT	CGGCACCCGA	TGCCTTAAAA	ATCATAACAG	AAAATGACCA	CGTCCTTCTC	6120
	TTCTGGAAAA	GTCTAGCTCT	AAAGGAAAAG	TATTTTAACG	AAAGCAGGGG	CTACGAGATA	618
	CACATGTTTG	ATAGCGCCAT	GAATATCACC	GCATACCTTG	GGAATACTAC	TGACAATTTC	624
45	TTTAAAATTT	CCAACCTGAA	GATGGGTCAC	AATTACACAT	TCACGGTCCA	GGCACGATGC	630
	CTTTTGGGCA	GCCAGATCTG	CGGGGAGCCT	GCCGTGCTAC	TGTATGATGA	GCTGGGGTCT	636
	GGTGGCGATG	CGTCGGCGAT	GCAGGCTGCC	AGGTCTACTG	ATGTCGCCGC	CGTGGTGGTG	642
50	CCCATCCTGT	TTCTGATACT	GCTGAGCCTG	GGGGTCGGGT	TTGCCATCCT	GTACACGAAG	648
	G1 T0001 000		ammas acaaa		CCCACTACAC	CTCCAGACTC	654

	GGCTCCGC	CA T	CTTC'	TCCT	C TG	GGGA	TGAC	TTG	GGGG.	AGG .	ATGA	TGAA	GA Ţ	GCTC	CTAT	G	۴600
	ATCACTGG	AT T	TTCG	GACG.	A CG	TCCC	CATG	GTG.	ATAG	cc							6639
5	(2) INFO	RMAT:	ION :	FOR	SEQ	ID N	0: 2	:									
10	(i)	(B)	UENC:) LEI) TY:) STI) TO:	NGTH PE: RAND	: 22 amin EDNE	13 a: o ac: SS:	mino id	S: aci	ds								•
	(ii)	MOLI	ECULI	E TY	PE:	prot	ein										
15	(xi)	SEQ	JENCI	E DE	SCRI	PTIO	N: SI	EQ II	on d	: 2:							
	Met 1	Ala	Thr	Arg	Ser 5	Ser	Arg	Arg	Glu	Ser 10	Arg	Leu	Pro	Phe	Leu 15	Phe	
20	Thr	Leu	Val	Ala 20	Leu	Leu	Pro	Pro	Gly 25	Ala	Leu	Cys	Glu	Val 30	Trp	Thr	
	Arg	Thr	Leu 35	His	Gly	Gly	Arg	Ala 40	Pro	Leu	Pro	Gln	Glu 45	Arg	Gly	Phe	
25		Val 50					55					60				_	
	65	Ala				70					75					80	
30		Ser			85					90					95		
	•	Leu		100					105					110			
35		Ser	115					120					125				
		Pro 130					135					140					
40	145	Asn				150					155					160	
		Ala			165					170					175		
45		Ile		180					185					190	•		
		Cys	195					200					205				
50		Leu 210					215					220					
	Ser 225	His	Pro	Asn	Lys	Gln 230	Leu	Trp	Lys	Ser	Asp 235	Asp	Phe	Gly	Gln	Thr 240	
55																	

	Trp	Ile	Met	Ile	Gln 245	Glu	His	Val	úys	Ser 250	₽he	Ser	Trp	Gly	Ile 255	qa.4
5	Pro	Tyr	Asp	Lys 260	Pro	Asn	Thr	Ile	Tyr 265	Ile	Glu	Arg	His	Glu 270	Pro	Ser
	Gly	Tyr	Ser 275	Thr	Val	Phe	Arg	Ser 280	Thr	Asp	Phe	Phe	Gln 285	Ser	Arg	Glu
10	Asn	Gln 290	Glu	Val	Ile	Leu	Glu 295	Glu	Val	Arg	Asp	Phe 300	Gln	Leu	Arg	Asp
	Lys	Tyr	Met	Phe	Ala	Thr 310	Lys	Val	Val	His	Leu 315	Leu	Gly	Ser	Pro	Leu 320
15	Gln	Ser	Ser	Val	Gln 325	Leu	Trp	Val	Ser	Phe 330	Gly	Arg	Lys	Pro	Met 335	Arg
	Ala	Ala	Gln	Phe 340	Val	Thr	Arg	His	Pro 345	Ile	Asn	Glu	Tyr	Tyr 350	Ile	Ala
20	Asp	Ala	Ser 355	Glu	Asp	Gln	Val	Phe 360	Val	Cys	Val	Ser	His 365	Ser	Asn	Asn
	Arg	Thr 370	Asn	Leu	Tyr	Ile	Ser 375	Glu	Ala	Glu	Gly	Leu 380	Lys	Phe	Ser	Leu
25	Ser 385	Leu	Glu	Asn	Val	Leu 390	Tyr	Tyr	Thr	Pro	Gly 395	Gly	Ala	Gly	Ser	Asp 400
	Thr	Leu	Val	Arg	Tyr 405	Phe	Ala	Asn	Glu	Pro 410	Phe	Ala	Asp	Phe	His 415	Arg
30	Val	Glu	Gly	Leu 420	Gln	Gly	Val	Tyr	Ile 425	Ala	Thr	Leu	Ile	Asn 430	Gly	Ser
	Met	Asn	Glu 435	Glu	Asn	Met	Arg	Ser 440	Val	Ile	Thr	Phe	Asp 445	Lys	Gly	Gly
35	Thr	Trp 450	Glu	Phe	Leu	Gln	Ala 455	Pro	Ala	Phe	Thr	Gly 460	Tyr	Gly	Glu	Lys
	Ile 465	Asn	Cys	Glu	Leu	Ser 470	Glu	Gly	Cys	Ser	Leu 475	His	Leu	Ala	Gln	Arg 480
40	Leu	Ser	Gln	Leu	Leu 485	Asn	Leu	Gln	Leu	Arg 490	Arg	Met	Pro	Ile	Leu 495	Ser
	Lys	Glu	Ser	Ala 500	Pro	Gly	Leu	Ile	Ile 505	Ala	Thr	Gly	Ser	Val 510	Gly	Lys
45	Asn	Leu	Ala 515	Ser	Lys	Thr	Asn	Val 520	Tyr	Ile	Ser	Ser	Ser 525	Ala	Gly	Ala
45	Arg	Trp 530	Arg	Glu	Ala	Leu	Pro 535	Gly	Pro	His	Tyr	Tyr 540	Thr	Trp	Gly	Asp
	His 545	Gly	Gly	Ile	Ile	Met 550	Ala	Ile	Ala	Gln	Gly 555	Met	Glu	Thr	Asn	Glu 560
50	Leu	Lys	Tyr	Ser	Thr 565	Asn	Glu	Gly	Glu	Thr 570	Trp	Lys	Ala	Phe	Thr 575	Phe

	Ser	Glu	Lys	Pro 580	Val	Phe	Val	Tyr	Gly 385	Leu	Leu	Thr	Glu	Pro 590	Gly	Giu
5	Lys	Ser	Thr 595	Val	Phe	Thr	Ile	Phe 600	Gly	Ser	Asn	Lys	Glu 605	Asn	Val	His
	Ser	Trp 610	Leu	Ile	Leu	Gln	Val 615	Asn	Ala	Thr	Asp	Ala 620	Leu	Gly	Val	Pro
10	Cys 625	Thr	Glu	Asn	Asp	Tyr 630	Lys	Leu	Trp	Ser	Pro 635	Ser	Asp	Glu	Arg	Gly 640
	Asn	Glu	Суѕ	Leu	Leu 645	Gly	His	Lys	Thr	Val 650	Phe	Lys	Arg	Arg	Thr 655	Pro
15	His	Ala	Thr	Cys 660	Phe	Asn	Gly	Glu	Asp 665	Phe	Asp	Arg	Pro	Val 670	Val	Val
	Ser	Asn	Cys 675	Ser	Cys	Thr	Arg	Glu 680	Asp	Tyr	Glu	Cys	Asp 685	Phe	Gly	Phe
20	Arg	Met 690	Ser	Glu	Asp	Leu	Ala 695	Leu	Glu	Val	Cys	Val 700	Pro	Asp	Pro	Gly
	Phe 705	Ser	Gly	Lys	Ser	Ser 710	Pro	Pro	Val	Pro	Cys 715	Pro	Val	Gly	Ser	Thr 720
25	Tyr	Arg	Arg	Ser	Arg 725	Gly	Tyr	Arg	Lys	Ile 730	Ser	Gly	Asp	Thr	Cys 735	Ser
	Gly	Gly	Asp	Val 740	Glu	Ala	Arg	Leu	Glu 745	Gly	Glu	Leu	Val	Pro 750	Cys	Pro
30	Leu	Ala	Glu 755	Glu	Asn	Glu	Phe	Ile 760	Leu	Tyr	Ala	Thr	Arg 765	Lys	Ser	Ile
	His	Arg 770	Tyr	Asp	Leu	Ala	Ser 775	Gly	Thr	Thr	Glu	Gln 780	Leu	Pro	Leu	Thr
3 <i>5</i>	Gly 785	Leu	Arg	Ala	Ala	Val 790	Ala	Leu	Asp	Phe	Asp 795	Tyr	Glu	His	Asn	Суs 800
	Leu	Tyr	Trp	Ser	Asp 805	Leu	Ala	Leu	Asp	Val 810	Ile	Gln	Arg	Leu	Cys 815	Leu
40	Asn	Gly	Ser	Thr 820	Gly	Gln	Glu	Val	Ile 825	Ile	Asn	Ser	Asp	Leu 830	Glu	Thr
••	Val	Glu	Ala 835	Leu	Ala	Phe	Glu	Pro 840	Leu	Ser	Gln	Leu	Leu 845	Tyr	Trp	Val
	Asp	Ala 850	Gly	Phe	Lys	Lys	Ile 855	Glu	Val	Ala	Asn	Pro 860	Asp	Gly	Asp	Phe
45	Arg 865	Leu	Thr	Val	Val	Asn 870	Ser	Ser	Val	Leu	Asp 875	Arg	Pro	Arg	Ala	Leu 880
	Val	Leu	Val	Pro	Gln 885	Glu	Gly	Ile	Met	Phe 890	Trp	Thr	Asp	Trp	Gly 895	Asp
50	Leu	ГÀЗ	Pro	Gly 900	Ile	Tyr	Arg	Ser	Asn 905	Met	Asp	Gly	Ser	Ala 910	Ala	Tyr
	Arg	Leu	Val	Ser	Glu	Asp	Val	Lys	Trp	Pro	Asn	Gly	Ile	Ser	Val	Asp

			915					920					925			
5	Asp	Gln 930	Trp	Ile	Tyr	Trp	Thr 935	Asp	Ala	Tyr	Leu	Asp 940	Сув	Ile	Glu	Arg
	Ile 945	Thr	Phe	Ser	Gly	Gln 950	Gln	Arg	Ser	Val	Ile 955	Leu	Asp	Arg	Leu	Pro 960
10	His	Pro	Tyr	Ala	Ile 965	Ala	Val	Phe	Lys	Asn 970	Glu	Ile	Tyr	Trp	Asp 975	Asp
70	Trp	Ser	Gln	Leu 980	Ser	Ile	Phe	Arg	Ala 985	Ser	ГÀЗ	Tyr	Ser	Gly 990		Gln
	Met	Glu	Ile 995	Leu	Ala	Ser	Gln	Leu 100	Thr O	Gly	Leu	Met	Asp 100		Lys	Ile
15	Phe	Tyr 1010	Lys)	Gly	Lys	Asn	Thr 101	Gly 5	Ser	Asn	Ala	Cys 102		Pro	Arg	Pro
	Cys 1025	Ser 5	Leu	Leu	Cys	Leu 1030	Pro	Arg	Ala	Asn	Asn 1035		Lys	Ser	Сув	Arg
20	Cys	Pro	Asp	Gly	Val 1045	Ala	Ser	Ser	Val	Leu 1050		Ser	Gly	Asp	Leu 1055	
	Cys	Asp	Cys	Pro 1060	Lys)	Gly	Tyr	Glu	Leu 1065		Asn	Asn	Thr	Cys 107	Val	Lys
25	Glu	Glu	Asp 1075	Thr	Cys	Leu	Arg	Asn 1080	Gln	Tyr	Arg	Cys	Ser 1089		Gly	Asn
	Cys	Ile 1090	Asn	Ser	Ile	Trp	Trp 1095	Сув	Asp	Phe	Asp	Asn 1100		Сув	Gly	qsA
30	Met 1105	Ser	Asp	Glu	Lys	Asn 1110	Cys)	Pro	Thr	Thr	Ile 1115		Asp	Leu	Asp	Thr 1120
	Gln	Phe	Arg	Cys	Gln 1125	Glu	Ser	Gly	Thr	Сув 1130		Pro	Leu	Ser	Tyr 1135	
35	Cys	Asp	Leu	Glu 1140	Asp	Asp	Cys	Gly	Asp 1145	Asn	Ser	Asp	Glu	Arg 1150	His	Cys
	Glu	Met	His 1155	Gln	Сув	Arg	Ser	Asp 1160	Glu	Tyr	Asn	Cys	Ser 1165		Gly	Met
40	Cys	Ile 1170	Arg	Ser	Ser	Trp	Val 1179	Cys	Asp	Gly	Asp	Asn 1180		Сув	Arg	Asp
•	Trp 1185	Ser	Asp	Glu	Ala	Asn 1190	Cys	Thr	Ala	Ile	Tyr 1195		Thr	Cys	Glu	Ala 1200
45	Ser	Asn	Phe	Gln	Су з 1205	Arg	Asn	Gly	His	Cys 1210		Pro	Gln	Arg	Trp 1215	
	Cys	Asp	Gly	Asp 1220	Ala	Asp	Cys	Gln	Asp 1225		Ser	Asp	Glu	Asp 1230	Pro	Ala
50	Asn	Cys	Glu 1235	Lys	Lys	Cys	Asn	Gly 1240	Phe	Arg	Сув	Pro	Asn 1245		Thr	Cys
	Ile	Pro 1250	Ser	Thr	Lys	His	Cys 1255		Gly	Leu	His	Asp 1260		Ser	Asp	Gly

	Ser 1265	Asp	Glu	Gln	His	Cys 1270		Pro	Leu	Cys-	Thr 1275		Phe	Met	Asp	Plie 1280
5	Val	Сув	Lys	Asn	Arg 1285		Gln	Cys	Leu	Phe 1290	His	Ser	Met	Val	Cys 1295	
	Gly	Ile	Ile	Gln 1300		Arg	Asp	Gly	Ser 1305		Glu	Asp	Pro	Ala 1310		Ala
10	Gly	Cys	Ser 1315		Asp	Pro	Glu	Phe 1320		Lys	Val	Cys	Asp 1325		Phe	Gly
	Phe	Gln 1330		Gln	Asn	Gly	Val 1335		Ile	Ser	Leu	Ile 1340		Lys	Cys	Asp
15	Gly 1345		Asp	Asp	Cys	Gly 1350		Tyr	Ser	Asp	Glu 1355		Asn	Cys	Glu	Asn 1360
	Pro	Thr	Glu	Ala	Pro 1365	Asn	Cys	Ser	Arg	Tyr 1370	Phe	Gln	Phe	Arg	Cys 1375	
20	Asn	Gly	His	Cys 1380		Pro	Asn	Arg	Trp 1385		Суз	Asp	Arg	Glu 1390		Asp
	Cys	Gly	Asp 1395		Ser	Asp	Glu	Lys 1400		Cys	Gly	Asp	Ser 1409		Val	Leu
25	Pro	Ser 1410		Thr	Pro	Ala	Pro 1415		Thr	Сув	Leu	Pro 1420		Tyr	Tyr	Arg
	Cys 1425		Gly	Gly	Ala	Cys 1430		Ile	Asp	Thr	Trp 1435		Cys	Asp	Gly	Tyr 1440
30	Arg	Asp	Cys	Ala	Asp 1445		Ser	Asp	Glu	Glu 1450	Ala	Сув	Pro	Ser	Leu 1455	
	Asn	Val	Thr	Ala 1460		Ser	Ser	Pro	Ser 1465		Pro	Gly	Arg	Cys 1470		Arg
	Phe	Glu	Phe 1475		Cys	His	Gln	Pro 1480	_	Lys	Cys	Ile	Pro 1485		Trp	Arg
35	Arg	Cys 1490		Gly	His	Gln	Asp 1495		Gln	Asp	Gly	Gln 1500		Glu	Ala	Asn
	Cys 1505		Thr	His	Ser	Thr 1510		Thr	Cys	Met	Ser 1515		Glu	Phe	Lys	Cys 1520
40	Glu	Asp	Gly	Glu	Ala 1525		Ile	Val	Leu	Ser 1530	Glu)	Arg	Cys	Asp	Gly 1535	
	Leu	Asp	Cys	Ser 1540		Glu	Ser	Asp	Glu 1549		Ala	Cys	Ser	Asp 1550		Leu
45	Thr	Val	Tyr 1555	-	Val	Gln	Asn	Leu 1560		Trp	Thr	Ala	Asp 1565		Ser	Gly
	Asn	Val 1570		Leu	Thr	Trp	Met 1575		Pro	Lys	Lys	Met 1580		Ser	Ala	Ala
50	Cys 1589		Tyr	Asn	Val	Tyr 1590		Arg	Val	Val	Gly 1595		Ser	Ile	Trp	Lys 1600

	Thr	Leu	Glu	Thr	His 1605		Asn	Lys	Thr	Asn 1613		Väl	Leu	Lys	Val 1619	
5	Lys	Pro	Asp	Thr 1620	Thr	Tyr	Gln	Val	Lys 1625		Gln	Val	Gln	Cys 1630		Ser
	Lys	Val	His 1635		Thr	Asn	Asp	Phe 1640		Thr	Leu	Arg	Thr 1645		Glu	Gly
10	Leu	Pro 1650		Ala	Pro	Gln	Asn 1655		Gln	Leu	Ser	Leu 1660		Gly	Glu	Glu
	Glu 1665	Gly	Val	Ile	Val	Gly 1670		Trp	Ser	Pro	Pro 1675		His	Thr	His	Gly 1680
15	Leu	Ile	Arg	Glu	Tyr 1685		Val	Glu	Tyr	Ser 1690		Ser	Gly	Ser	Lys 1695	
	Trp	Thr	Ser	Glu 1700	Arg	Ala	Ala	Ser	A sn 1705		Thr	Glu	Ile	Lys 1710		Leu
20	Leu	Val	Asn 1715		Leu	Tyr	Thr	Val 1720		Val	Ala	Ala	Val 1725		Ser	Arg
	Gly	11e 1730		Asn	Trp	Ser	Asp 1735		Lys	Ser	Ile	Thr 1740		Val	Lys	Gly
25	Lys 1745		Ile	Pro	Pro	Pro 1750		Ile	His	Ile	Asp 1755		Tyr	Asp	Glu	Asn 1760
	Ser	Leu	Ser	Phe	Thr 1765		Thr	Val	Asp	Gly 1770		Ile	Lys	Val	Asn 1775	
3 <i>0</i>	Tyr	Val	Val	Asn 1780	Leu)	Phe	Trp	Ala	Phe 1785		Thr	His	Lys	Gln 1790		Lys
	Lys	Thr	Met 1795		Phe	Gln	Gly	Ser 1800		Val	Ser	His	Lys 1805		Gly	Asn
3 <i>5</i>	Leu	Thr 1810		Gln	Thr	Ala	Tyr 1815		Ile	Ser	Ala	Trp 1820		Lys	Thr	Asp
	Leu 1825		Asp	Ser	Pro	Leu 1830		Phe	Glu	His	Val 1835		Thr	Arg	Gly	Val 1840
	Arg	Pro	Pro	Ala	Pro 1845		Leu	Lys	Ala	Arg 1850		Ile	Asn	Gln	Thr 1855	
40	Val	Glu	Cys	Thr 1860	Trp)	Thr	Gly	Pro	Arg 1865		Val	Val	Tyr	Gly 1870		Phe
	Tyr	Ala	Thr 1875		Phe	Leu	Asp	Leu 1880		Arg	Asn	Pro	Ser 1885		Leu	Thr
45	Thr	Pro 1890		His	Asn	Ala	Thr 1895		Leu	Val	Gly	Lys 1900		Glu	Gln	Tyr
	Leu 1905		Leu	Val	Arg	Val 1910		Met	Pro	Tyr	Gln 1915		Pro	Ser	Ser	Asp 1920
50	Tyr	Val	Val	Val	Lys 1925		Ile	Pro	Asp	Ser 1930	_	Leu	Pro	Pro	Arg 1935	
	Leu	His	Ala	Val	His	Thr	Gly	Lys	Thr	Ser	Ala	Val	Ile	Lys	Trp	Glu

		1940	2045		
			1943		1950
5	Ser Pro Tyr 195	Asp Ser Pro A	Asp Gln Asp Let 1960	ı Phe Tyr Ala 196	Ile Ala Val
	Lys Asp Leu 1970	Ile Arg Lys T	Chr Asp Arg Sei .975	Tyr Lys Val 1980	Lys Ser Arg
10	Asn Ser Thr 1985	Val Glu Tyr T 1990	Thr Leu Ser Lys	Leu Glu Pro 1995	Gly Gly Lys 2000
	Tyr His Val	Ile Val Gln L 2005	eu Gly Asn Met 201	Ser Lys Asp .0	Ala Ser Val 2015
15	Lys Ile Thr	Thr Val Ser L 2020	eu Ser Ala Pro 2025	Asp Ala Leu	Lys Ile Ile 2030
,,,	Thr Glu Asn 2035	Asp His Val L	eu Leu Phe Trp 2040	Lys Ser Leu 204!	
	Glu Lys Tyr 2050	Phe Asn Glu S	er Arg Gly Tyr 055	Glu Ile His 2060	Met Phe Asp
20	Ser Ala Met 2065	Asn Ile Thr A 2070	la Tyr Leu Gly	Asn Thr Thr 2075	Asp Asn Phe 2080
	Phe Lys Ile	Ser Asn Leu L 2085	ys Met Gly His 209		Phe Thr Val 2095
25	Gln Ala Arg	Cys Leu Leu G 2100	ly Ser Gln Ile 2105	: Cys Gly Glu	Pro Ala Val 2110
	Leu Leu Tyr 2115	Asp Glu Leu G	ly Ser Gly Gly 2120	Asp Ala Ser 2125	
30	Ala Ala Arg 2130	Ser Thr Asp V	al Ala Ala Val 135	Val Val Pro 2140	Ile Leu Phe
· .	Leu Ile Leu 2145	Leu Ser Leu G 2150	ly Val Gly Phe	Ala Ile Leu 2155	Tyr Thr Lys 2160
35	His Arg Arg	Leu Gln Ser Se 2165	er Phe Thr Ala 217		Ser His Tyr 2175
	Ser Ser Arg	Leu Gly Ser A 2180	la Ile Phe Ser 2185	Ser Gly Asp	Asp Leu Gly 2190
40	Glu Asp Asp 2195	Glu Asp Ala P	ro Met Ile Thr 2200	Gly Phe Ser 2205	
	Pro Met Val 2210	Ile Ala			
45 (2)	INFORMATION F	OR SEQ ID NO:	3:		
•	(i) SEQUENCE (A) LEN (B) TYP (C) STR	CHARACTERIST GTH: 6961 base E: nucleic ac ANDEDNESS: do	ICS: e pairs id uble		
50		OLOGY: linear			

(ix)	FEAT	JRE:	
	(A)	NAME/KEY: sig	peptide
	(B)	LOCATION: 178.	.261

5

(ix) FEATURE:
 (A) NAME/KEY: mat peptide
 (B) LOCATION:262..6816

		(xi)	SEO	UENCI	E DES	CRI	PTIO	N: SI	EQ II	D NO	: 3:						
10	aaaa	77.00		አ ሮ አ ሮ (ግጥርን ሲ	- GG	רנירר	GCGC	CGC	GCCG	CGC	CGCG GCCC	CCGA CGCC	GC G GC T	GGAC	CCAGC CGGCG	60 120
	GCGC	3CTG	CA C	ATTC"	TCTC	TG	GCGG	CGGC	GCC	ACCT	GCA	GCCG	CGTT	CG C	CCGA	ACATG Met 1	180
15	GCG Ala	ACA (CGG Arg	AGC A	AGC A	AGG Arg	AGG Arg	GAG '	Ser	CGA Arg	CTC Leu	CCC Pro	TTC Phe	CTA Leu 15	TTC Phe	ACC Thr	228
	CTG Leu	GTC Val	Ala	CTG Leu	CTG Leu	CCG Pro	CCC Pro	Gly .	10 GCT Ala	CTC Leu	TGC Cys	GAG Glu	GTG Val	TGG	ACG Thr	CGG Arg	276
20	ACA Thr	CTG Leu	20 CAC His	GGC Gly	GGC Gly	CGC Arg	Ala	25 CCC Pro	TTA Leu	CCC Pro	CAG Gln	Glu	CGG	GGC Gly	TTC Phe	CGC Arg	324
	GTG Val	35 GTG Val	CAG Gln	GGC Gly	GAC Asp	CCG Pro	40 CGC Arg	GAG Glu	CTG Leu	CGG Arg	Leu	45 TGG Trp	GAG Glu	CGC A rg	GGG Gly	Asp	372
25	50	3.00	ccc	CCG	»GC	55 CGG	GCG	GAC	GAG	AAG	CCG	CTC	CGG	AGG	AGA Arg	CGG	420
	100	CI CITT	ccc	CTC	70 CAG	כככ	GAG	CCC	ATC	75 AAG	GTG	TAC	GGA	CAG	80 GTC Val	AGC	468
30	CITICS.	አለጥ	CNT	85 TCC	CAC	דממ	CAG	ATG	90 GTG	GTG	CAC	TGG	GCC	GGA	GAG Glu	AAA	516
	3.00	7 7 C	100	አጥር	GTG	GCC	ፐፕር	GCC	CGG	GAC	AGC	CTG	GCG	TTG	GCC Ala	AGG	564
	505	115	300	N CUTT	ርእጥ	CTC	120	GTG	тст	ТАТ	GAC	TAT	GGA	AAA	TCA	TTC	612
35	130	220	እጥጥ	тсъ	GAG	135	TTG	AAC	TTC	GGC	GCG	GGA	AAT	AAC	Ser ACA	GAG	660
	Asn	Lys	Ile	Ser	Glu 150	TTC	Leu	CAC	Phe	155 CCT	GCG	GAC	ASN	AAA!	160 CGG	TAC	708
40	Ala	Val	Val	Ala 165	Gln	Phe	Tyr	His CAG	Ser 170	Pro CTC	TGG	ASP ATC	AST : ACG	175 177	GAC	TTC	756
	Ile	Phe	Ala	Asp	Ala	Tyr	Ala	. Gln 185	Tyr	Leu	Trp	116	190)	: Asp	Phe	804
45	Cys	Asn	Thr	Ile	His	Gly	Phe	Ser	Ile	Pro	Phe	205	i Alg	l Alc	r war) Leu	852
	Leu	Leu	His	Ser	Lys	Ala	Ser	Asn	Leu	Lev	220) 1 GT)	Pile	. Apl	פותיי	S TCT Ser 225	900
50	His	Pro	Asn	Lys	Gln	Leu	Tr	Lys	Ser	C AST	ASI	Pne	3 613	, GI	240		
	ATC Ile	ATG Met	ATI : Ile	CAA Gln 245	ı Glu	CAC His	C GTC	AAG Lys	Ser 250	r Phe	r TCT e Sei	r TG(e GGI	25:	e wat	Pro	948

			26	n				- 191	. 11	= (-T	u Ar	g ni	Ω G.⊤	u Pr	o Se	T GGC	
5	TA:	C TC r Se 27	C AC r Th: 5	G GT	r TTC	CGA Arg	AGT Ser	ACA	A GA	C TTO Pho	C TT	C CA	G TC n Se	C CC r Ar	G GI	A AAC u Asn	1044
	290	0				295	GAA Glu	GTG Val	. AL	y AS	o Pue	e GI;	G CT	u Ar	g As	C AAG	
10	•				310	- -	val	val	. nis	Jet 316	TTO Let	G GGG	/ Se:	r Pr	o Le	305 G CAG u Gln	1140
				325			Val	Ser	220	; GT	Arc	j Lys	Pro	o Me	t Ar	0 G GCC g Ala	1188
15			340)		AL 9	nis	345	TIE	: ASI	ı GIV	тул	Туз	C AT	C GC e Al	G GAT a Asp	1236
		355	,	op		VUI	360	Val	Cys	vaı	Ser	His	Ser	As	n As	C CGC n Arg	1284
20	370				110	375	GIU	Ald	GIU	СТУ	Leu	Lys	Phe	Se:	r Le	G TCC u Ser 385	1332
					390	+ y -	TYL	1 111	PLO	GTA	GIY	Ala	Gly	Sei	: As	385 C ACC D Thr	1380
25			9	405	1110	ALA	MSII	GIU	430	Pne	Ala	Asp	Phe	His	Ar	o F GTG J Val	1428
		2	420		013	Val	IYL	116	АТА	Inr	Leu	Ile	Asn	Gl	' Sei	ATG Met	1476
30	TGG	435 GAA	TTT	CTG	CAG	GC.A.	440 CCA	CCC	TIE	rnr	Phe	Asp 445	Гув	Gly	Gl _}	Thr	1524
	450 AAC	TGT	GAG	CTG	TCC	455 GAG	GGC	ATA	TCC	Thr	Gly 460	Tyr	Gly	Glu	Lys	1le 465	1572
35		-1-			470	Giu	GIÀ	cys	ser	Leu	His	Leu	Ala	Gln	Arg	Leu	1620
	GAG	TCG	GCG	485 CCT	GGC	CTC	OIII Δጥሮ	ren	490 CCC	Arg	Met	Pro	Ile	Leu 495	Ser	Lys	1668
40	TTG	GCT	500 AGC	AAG	ACA	AAC (TTG	505 TAC	ATC	TOT	GIY	Ser	Val 510	Gly	Lys	Asn	1716
	TGG	515 CGA	GAG	GCA	СТТ	፲ ፫፻፲ (520	C.C.II.	TIE	ser	Ser	525	Ala	Gly	Ala	Arg	1764
45	530 GGC	GGC	ATC	ATC	ATG (535 GCC <i>i</i>	SIY . ATT (GCC	CAA	ccc	Tyr 540	Thr	Trp	Gly	Asp	His 545	1812
	AAG	TAC	AGT	ACC .	550 AAC (GAA (ace (SAC	ACC ACC	555 TCC	Met	GIU	Thr	Asn	Glu 560	Leu	1860
50	GAG	AAG	ccc	565 GTG	TTT (STG 1	י עבר	31u 366 /	570	Trp	гуs	Ala	Phe	Thr 575	Phe	Ser	1908
	Glu AGC	-,,	580	vaı	FILE	val 1	yr	31Y 1	Leu .	Leu	Thr	Glu	Pro	Gly	Glu	Lys	1956 2004
												-					2004

	Ser		Val	Phe	Thr		Phe	Gly	Ser	Asr.	Lys	Glu 605	Asn	Val	His	ser	
	TGG	595 CTC	ATC	CTC	CAG	GTC :	דבב	GCC	ACA	GAC	GCC	CTG Leu	GGG Glv	GTT Val	CCT Pro	TGC Cvs	2052
5	610					615					620					625	
	3.03	GAG	AAC	GAC	ТЪС	AAG	CTC	TGG	TCA	CCA	TCT	GAT	GAG	CGG	GGG	TAA	2100
	Thr	Glu	Asn	Asp	Tyr	Lys	Leu	Trp	Ser	Pro 635	Ser	Asp	Glu	Arg	Gly 640	Asn	
	a.c	ጥረጥ	TTC	بلملت	630 GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	AGG	ACC	CCG	CAC	2148
	GAG	Cvs	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro	His	
10		_		CAE					650					655			2196
	GCC	ACA	TGC	TTT	AAC	GGA	GAA	GAC	TTT	GAC	AGG	CCG Pro	Val	Val	Val	Ser	2190
			cco					665					6/0				
	AAC	TGC	TCC	TGC	ACC	CGG	GAG	GAC	TAT	GAG	TGT	GAC	TTT	GGC	TTC	CGG	2244
	Asn	Cys	Ser	Суз	Thr	Arg	Glu	Asp	Tyr	Glu	Cys	Asp	Phe	Gly	Phe	Arg	
15		C75					680					685 CCA					2292
	ATG	AGT	GAA	AGD	Len	Ala	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Gly	Phe	
						695					700					, , ,	
	~~~	GGA	AAG	TCC	TCC	CCT	CCA	GTG	CCT	TGT	CCC	GTG	GGC	TCT	ACG	TAC	2340
	Ser	Gly	Lys	Ser		Pro	Pro	Vai	Pro	715	Pro	Val	GIY	261	720	1 7 1	
20	NCC.	CCA	TCA	AGA	710 GGC	TAC	CGG	AAG	ATT	TCT	GGG	GAC	ACC	TGT	AGT	GGA	2388
	Ara	Arg	Ser	Arq	Gly	Tyr	Arg	Lys	Ile	Ser	Gly	Asp	Thr	CAP	Ser	Gly	
	_			7つに					7.30					133			2436
	GGA	GAT	GTT	GAG	GCA	CGG	CTA	GAA	GGA	GAG	Ten	GTC Val	Pro	Cvs	Pro	Leu	2430
			740					745					/50				
25	GCA	GAA	CNG	AAC	GAG	TTC	ATC	CTG	TAC	GCC	ACC	CGC	AAG	TCC	ATC	CAC	2484
	Ala	Glu	Glu	Asn	Glu	Phe	Ile	Leu	Tyr	Ala	Thr	Arg 765	Lys	ser	TTE	HIS	
		755	~~~	- CEC	CCT	TOO	760	»CC	ACG	GAC	CAG	TTG	CCC	CTC	ACT	GGG	2532
	CGC	TAL	Agn	T.eu	Ala	Ser	Glv	Thr	Thr	Glu	Glr	Leu	Pro	Leu	Thr	Gly	
	770					775					701	,				, 0 3	2500
30	mmo		GCA	GCA	GTG	GCC	CTG	GAC	TTI	GA(	TAT	GAG	CAC	AAC	TGC	CTG	2580
	Leu	Arg	, Ala	Ala	Val	Ala	Leu	Asp	Pne	79!		GIU	nis	ASI.	800	Leu	
	ייאייי	TCC	י יייריים	י מאר	790 יייייייייייייייייייייייייייייייייייי	CCC	TTG	GAC	GTC	AT	CAC	CGC	CTC	TGT	TTC	AAC	2628
	Tvi	Tr	Ser	Asp	Leu	Ala	Leu	Asp	Val	. Ile	e Gli	n Arg	Leu	. Cys	. ner	Asn	
				0 0 5					810	3				01-	,		2676
35	GGC	AGT	ACA	A GGA	CAA	GAG	GTG	ATC	: ATC	AA!	n Se	r Ast	Leu	Gli	Thi	GTA Val	20.0
			~~	•				ロソト					0,50	,			
	GAA	A GC	n mm/	7 00	r TTI	GAA	ccc	CTC	AG	CA	A TT	A CTT	TAC	TGO	GTO	GAC Asp	2724
	Glu	ı Ala	a Le	ı Ala	a Phe	Glu	Pro	Let	ı Sei	c Gl	n Le	u Lei 845	TAT	Tr	o va.	L Asp	
	~~	83	5			י איייר	840 GAG	) ፤ ርጥጀ	4 60	~ AA	т сс	A GAT	GG7	GA	C TT	CGA	2772
40	7 G C	A GG	v Ph	o Turs	LVS	Ile	Glu	val	Ala	a As	n Pr	o Ası	Gly	/ As	p Pho		
		_				055					- 00	U					2820
	- m		CGT	CGT	C AA	r TCC	TCC	GTO	G CT	G GA	T CG	G CC	CGC	TAI	a Le	G GTC	2020
	Le	u Th	r Va	l Va	l Ası	n Ser	Sei	c va.	r ne.	u AS 87	у <u>ж.</u> 5	g rr	JAL	,	88	0	
	CT	ጥ ርጥ	c cc	C CA	870 A GA	N CCC	ATO	CATO	G TT	C TG	G AC	C GA	C TG	G GG	A GA	C CTG	2868
45	Le	u Va	l Pr	o G1:	n Gl	u Gly	/ Ile	e Met	t Pn	e II	p Th	r As	o Tr	5 67	y no	p Leu	
				00	_				ж ч	()				0,2	_		
	AA	G CC	T GG	G AT	T TA	T CGC	AG(	AA!	c AT n Me	t As	n Gl	y Se	r Al	a Al	a Ty	T CGC r Arg	3
			~ ~	^				40	<b>-</b>				7 =	•			
	CT	C GI	c mc	~ ~ ~	G GA	T GT	S AA	G TG	G CC	C AF	AT GO	C AT	T TC	CGT	G GA	C GAT	2964
50	Le	u Va	l Se	r Gl	u As	p Va	l Ly	s Tr	p Pr	O As	sn Gl	y II. 92	e 3e	r va	T AS	p Asp	
	~	91	.5	ירי די	C TC	G AC	92 2 CA	ሀ ጥ GC	с та	c c	rg ga	C TG	C AT	T GA	G CG	C ATC	3012
	G1	n Ti	no Il	e Tv	r Tr	p Th	r As	p Al	a Ty	r Le	eu As	эр Су	s Il	e Gl	u Ar	g Ile	:
				•		-											

	930 935 ACG TTC AGC GGC CAG CAG CGC TCC GTC AFC CTG GAC AGA CTC CCG CAC Thr Phe Ser Gly Gln Gln Arg Ser Val IIe Leu Asp Arg Leu Pro His	)
5	CCC TAT GCC ATT GCT GTC TTT ANG AND GAG ATT GCC 960	,
	965 970 975 TCA CAG CTC AGC ATA TTC CGA CCT TCT AND TTC PAGE 175	
10	980 And Ser Lys lyr Ser Gly Ser Gln Met	
	GAG ATT CTG GCC AGC CAG CTC ACG GGG CTG ATG GAC ATG AAG ATC TTC 3204 Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile Phe 1005	
15	TAC AAG GGG AAG AAC ACA GGA AGC AAT GCG TGT GTA CCC AGG CCG TGC 3252 Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro Cys 1010 1020 1025	
75	Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg Cys	
	CCA GAT GGC GTG GCC AGC AGT GTC CTC CCT TCC GGG GAC CTG ATG TGT Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met Cvs	
20	GAC TGC CCT AAG GGC TAC GAG CTG AAC AAC ACG TGT GTC AAA GAA 3396	
	1060 1065 Ash Ash Thr Cys Val Lys Glu	
	GAA GAC ACC TGT CTG CGC AAC CAG TAC CGC TGC AGC AAC GGG AAC TGC Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn Cys 1075	
25	Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp Met	
	AGC GAC GAG AAG AAC TGC CCT ACC ACC ATC TGC GAC CTG GAC ACC CAG 3540 Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr Gln	
30	Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys Cys	
	Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys Glu	
35	Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met Cys	
	Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp Trp	
40	Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala Ser	
	ARC TTC CAG TGC CGC AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG TGT 3828 Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala Cys 1205	
45	Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala Asn	
	TGT GAG AAG AAG TGC AAC GGC TTC CGC TGC CCG AAC GGC ACC TGC ATT 3924  Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys Ile	
50	Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly Ser	
	GAC GAG CAC TGC GAG CCC CTG TGT ACA CGG TTC ATG GAC TTC GTG 4020 Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe Val 1270 1280	

	Сув	Lys	Asn	CGC Arg 1285	Gln	Gln	Cys	Leu	Phe 1290	his	ser	Met	۷ĕ٦	Cys 1295	Asp	GIÀ	4063
5	ATC Ile	ATC Ile	CNG	TGC Cys	CCT	GAC Asp	GGC Gly	TCC	GAC	GAG	GAC Asp	CCA Pro	GCC Ala	TTT Phe	GCA Ala	GGA Gly	4116
			1200	١				1305					13T0	,			
	TGC	TCC	CCA	CAC	CCC	GAG	TTC	CAC	AAG	GTG	TGC	GAT	GAG	TTC	GGC	TTC	4164
	Cys	Ser	Arg	Asp	Pro	Glu	Phe	His	Lys	Val	Cys	Asp	GIU	Phe	Gly	Phe	
		1215					1320	)				1323	)				40.0
	CAG	TGT	CAG	AAC	GGC	GTG	TGC	ATC	AGC	TTG	ATC	TGG	AAG	TGC	GAC	GGG	4212
10	Gln	Cys	Gln	Asn	Gly	Val	Cys	Ile	Ser	Leu	Iте	Trp	Lys	Cys	Asp	GIA	
						1225					1340	)				T343	4360
	ATG	GAT	GAC	TGC	GGG	GAC	TAC	TCC	GAC	GAG	GCC	AAC	TGT	GAA	AAC	000	4260
	Met	Asp	Asp	Cys	Gly	Asp	Tyr	Ser	Asp	Glu	Ala	Asn	Сув	GIU	ASII	PIO	
					1350	)				1355	22.0	mm.c	000	TOT	1360		4308
	ACA	GAA	GCC	CCC	AAC	TGC	TCC	CGC	TAC	TTC	CAG	110	7~4	CAC	JAC	Acn	4500
15	Thr	Glu	Ala	Pro		Cys	Ser	Arg	Tyr	Pne	GIn	Pne	Arg	137	msp :	ASII	
				1365	5				1370	, mcm	CAC	N.C.C	CAG			тст	4356
	GGC	CAC	TGC	ATC	CCC	AAC	AGG	TGG	AAG	Corn	ACD	AGG	GAG	Acn	Aen	Cvs	4330
	Gly	His		Ile	Pro	Asn	Arg	110	гуя	Cys	ASP	rra	1390	7	rup	0,0	
			1380	0	~~~	a.a	330	1385	TOTAL D	CCA	CAT	тсъ			CTT	CCG	4404
	GGG	GAC	TGG	TCC	GAC	GAG	AAG	3 am	191	Clv	Acn	Car	Hig	Val	Leu	Pro	
20	Gly			Ser	Asp	GIU	Lys	Asp	Cys	Gry	ASP	140	5				
		139	5	CCT	CC3	ccc	140	ארכב	тст	CTG	CCC	AAT.	TAC	TAC	CGC	TGC	4452
	TCT	ACG	ACT	Pro	GCA 3.1-	CCC	Sor	Thr	Cve	Len	Pro	Asn	Tvr	Tvr	Arg	Cys	
			Thr	Pro	Ald	1419	367	1111	Cys		142	0	-1-	-1-		1425	
	1410	, , , ,	~~~	GCC	TCC	CTC	מדמ מדמ	CAC	ACG	TGG	GTT	TGT	GAC	GGG	TAC	CGA	4500
	GGC	GGG	GGG	Ala	Cva	Val	Tle	Agn	Thr	Tro	Val	Cvs	Asp	Gly	Tyr	Arg	
25	GIĀ	GIY	GIY	Ala	143	Nai	116	ASP	****	143	5	-7-	E	2	144	0	
	a.m	maa	CCN	GAT	CCA	TCC	GAC	GAG	GAA	GCC	TGC	CCC	TCG	CTC	CCC	AAT	4548
	GAT	TGC	λla	Agn	Glv	Ser	Asp	Glu	Glu	Ala	Cvs	Pro	Ser	Leu	Pro	Asn	
	_			1 4 4	_				145	U				140	<b>ગ</b>		
	CTC	አ ርጥ	GCC	ACC	TCC	TCC	CCC	TCC	CAG	CCT	GGA	CGA	TGC	GAC	CGA	TTT	4596
	7721	Thr	Δla	Thr	Ser	Ser	Pro	Ser	Gln	Pro	Gly	Arg	Сув	Asp	Arg	Phe	
30			716	^				146	5				14/	U			
	GAG	ጥጥጥ	CAC	TOO	CAC	CAG	CCA	AAG	AAG	TGC	ATC	CCT	AAC	TGG	AGA	CGC	4644
	Glu	Phe	Glu	Cvs	His	Gln	Pro	Lys	Lys	Cys	Ile	Pro	Asn	Trp	Arg	Arg	
		4 77	-				148	O.				140					
	TGT	CAC		CAT	CAG	GAT	TGC	CAG	GAT	' GGC	CAG	GAC	GAC	GCC	: AAC	TGC	4692
	Cvs	Asp	Glv	His	Gln	Asp	Cys	Gln	Asp	Gly	r Gln	Asp	) Glu	Ala	Asr	LCYS	
35	4 4 0	^				149	5				150	טי				1000	4740
-	000	200	CAC	: AGC	ACC	TTG	ACC	TGC	ATO	AGC	TGG	GAC	TTC	AAC	TGI	GAG	4740
	Pro	Thr	His	Ser	Thr	Leu	Thr	: Cys	Met	: Ser	cirr	) Gli	ı Phe	: гу	, су	, Gra	
					7 5 7	Λ.				151	7				104		4788
	GAT	' GGC	GAG	GCC	TGC	: ATC	GTC	CTC	TC	A GAA	A CGC	1 TGC	CAU		. Dhe	CTG	4700
	Asp	Gly	/ Glu	ı Ala	ι Суз	: Ile	: Val	L Leu	ı Sex	c GIV	1 Arg	j Cys	s Asi	נים ע	FILE	e Leu	
40				152	) <b>C</b>				15.	30				10.	, ,		4836
40	GAC	TGC	TCF	A GAT	GAC	AGC	GAC	GAG	AAC	3 GCC	TGC	. AG	r GA.	Cl	, T.O.	A ACT	1050
	Asp	Cys	s Sei	c Asp	o Glu	ı Ser	: Asr	GIU	r PA	s Ala	а су	s se	15!	: V	י אכי	ı Thr	
			154	10				154	15	3 80		r ca			r GG	TAA F	4884
	GTA	A TAC	CAA	A GT	A CAC	AA7	r CT.	r CAC	TG	ACA	A GC.	I GAV	o Dha	- FO	r G1	G AAT	
	Va]	L Ty	r Lya	s Val	l Gli	n Ası	ı Let	ı Gli	i irl	p Thi	r Ale	15	5 FIN	- JC.	L 01.	y Asn	
45		159	55				150	50	~ ~ ~ ~		N NT/	2 CC	ר שר.	ד מכי	r GC	r TGT	4932
45	GT	AC'	r TTC	G AC	C TGO	3 ATC	اخان خ	- CCC	AA	A AA	a Me	b Dr	o Se	r Al	a Al	T TGT	
			r Le	u Thi	r Tr	o Met	. Arg	g Pro	) PA	э гу	15	80	<b>.</b>			a Cys 1585	5
	15	70	_			15	/5 		n (111	T CC	N (2)	G AG	с ат	A TG	GAA		4980
	GT	A TA	C AA	C GT(	J TA	L IA	I AG	m UI.	ı Gl	ינט ג רבי ן	י מו	ים אינו	r II	e Tr	ם Lv	G ACT s Thr	
	Va.	ту:	r As	n va	TIY	E TÄ:	LAF	y va.	r va	15	7 GI				16	s Thr 00	
50	,-up-		~ .~	m ~~	15:	ጋ∪ ሮ አኦነ	אמ ד	G AC	44 4	C AC	T GT	A TT	A AA	A GT		G AAA	5028
50	CT	GA!	G AC	L CA	CAG	r ye	n Tar	g Th	r λe	n Th	r Va	l Le	u Lv	s Va	l Le	u Lys	
				96	ΛE				16	10				70			
	CC	אבי ע	ጥ አር	L 7C 70	ር ጥል C	C CA	G GT	T AA	A GT	G CA	G GT	T CA	G TG	CCI	G AG	C AAG	5076
		_ (74						~									

	Pro	Ası	7h	r Th	r Ty	r Gl	n Va	l Ly	s Va	l Gi	n Va	1 G1	n Cy	′∈ I-e	u Se	r Lys	
	GTG	CAC	CAA	CAC	יממ "	ጥ ርአ		T 0	~ . ~	~		_	Τ 6	30			•
5	Val	His	As:	n Th	r As	n As		e va	G ACC	r Le	G AG u Ar	g in	r Pr	A GA O Gl	G GG u Gl	A TTG	5124
	CCA	GAC	GC	C CC	T CA	3 22	سي م	~ ~ ~ ~ ~ ~ ~	a ama	, ma		16	45				
	Pro	Asp	Al	a Pr	o Gl	n Ası	n Le	n Gl	n Lei	, SO	G CT	C CA	C GG	G GA	A GA	G GAA u Glu	5172
	165	0				16	55	u G1.	n Dec	, se.	r Le	u Hi	a GI	y Gl	u Gl	u Glu	
	GGT	GTG	AT'	T GT	G GG	CAC	~ TC	G AG	ר ככיי	ר ככי	16:	~ ~-	a			1665 C CTC	5
	Gly	Val	. Ile	e Va	l Gly	/ His	Tr	o Ser	r Pro	Pr	o Th	c CA	C AC	C CA	C GG	C CTC	5220
10					16	70				16	75 75	r ur	s in	r Hi	s GI	y Leu	
	ATT	CGC	GA	ATA	C AT	GT!	A GAG	3 TAT	r ago	300	~ 3.00	rcc	T TO		16	80 G TGG	
	Ile	Arg	Gli	ı Ty:	r Ile	≥ Val	Gli	ı Tyr	r Ser	Arc	Sei	r G1	1 IC	C AA	GT	G TGG l Trp	5268
				16	85			-	169	0	,	- 01	y se.	160	s va.	ı Trp	
	ACT	TCA	GAZ	A AG	G GC1	C GC1	' AG	C AAC	TTI	' AC	A GAZ	A AT	A AA	S AAC	יידוייתייי ככ	G TTG	F316
	Thr	Ser			g Ala	Ala	Sez	Asr	ı Phe	Th	Gli	ı Ile	e Lv	S Acr	- II.	G TTG	5316
15																	
	GTC Val	AAC	ACC	CTC	3 TAC	ACC	GTC	AGA	A GTG	GCT	r GCC	GT(		~	r cgr	r ggg	5364
	vai	171	THI	Let	л Туг	Thr	val	- wrā	j Val	Ala	a Ala	a Vai	L Thi	r Sei	Arc	r GGG	2364
	Tle	GL	AAC	160	AGC	GAI	TCC	: AAA	TCC	ATT	ACC	AC	GTO	G AAA	GG2	A AAA	5412
	1730	Ory	ASI	TIE	Ser	Top	, ser	. Lys	Ser	Ile	Thr	Thi	· Val	l Lys	Gly	A AAA / Lys	
20																	
	Ala	Tle	Pro	Dre	Dro	WAI	AIC	CAC	ATT	GAC	AAC	TAC	GAT	GAA	LAA 1	1745 TCC	5460
				110	175	HOIL	ııe	HIS	ıre	Asp	) Asn	туз	Asp	Glu	Asr	TCC Ser	
	CTG	AGT	TTT	ACC	TOTO	്മന	GTG	CAT		175	5				17€	O TAT	
	Leu	Ser	Phe	Thr	Leu	Thr	Val	) CAI	C1	AAC	ATC	AAC	GTC	AAT	, GGC	TAT Tyr	5508
0.5																	
25	GTG	GTG	AAC	CTT	TTC	TGG	GCA	ماست	CAC	ACC.	C 2 C	* **	<i>-</i>	177		AAA	
	Val	Val	Asn	Leu	Phe	Trp	Ala	Phe	Asn	Thr	Lic	AAA	CAA	GAG	AAG	AAA Lys	5556
			178	0				178	5 5	****	nis	гуу	170	r GTI	råa	Lys	
	ACC	ATG	AAC	TTC	CAA	GGG	AGC	TCA	GTG	TCC	CAC	מממ	179		3 3 7	CTG	
				Phe	Gln	Gly	Ser	Ser	Val	Ser	His	Tive	. Ull	C1	AAT	CTG Leu	5604
00																	
30	ACA Thr	GCA	CAG	ACG	GCC	TAT	GAG	ATT	TCC	GCC	TGG	~~~		ልርጥ	GAC	TITLE	5.55
		Ala	Gln	Thr	Ala	- 1 -	GIU	Ile	Ser	Ala	Trp	Ala	Lvs	Thr	Agn	Lau	5652
	GGC Glv	GAT.	AGT	CCT	CTG	TCA	TTT	GAG	CAT	GTC	ACG	ACC	AGA	GGG	GTT	CGC	5700
	Gly .	ASP	ser	Pro	<u> Leu</u>	Set	Phe	Glu	His	Val	Thr	Thr	Arg	Gly	Val	Ara	3,00
35																	
35	CCA Pro	Pro	Al=	Dro	AGC	CTC	AAG	GCC	AGG	GCT	ATC	AAT	CAG	ACT	GCA	GTG	5748
	Pro			184	261	Leu	гÀз	Ala	Arg	ALA	Ile	Asn	Gln	Thr	Ala	Val	
	GAA (	Cys	Thr	Trp	Thr	Glv	Pro	AGG A×a	AAI	GIG	GIG	TAT	GGC	ATT	TTC	TAT	5796
40	GCC A	ACA	TCC	TTC	CTG	GAC	CTC	TAC	CGC	מממ	CCA	300	1870	~~~			
***	Ala 1	Thr	Ser	Phe	Leu	Asp	Leu	Tvr	Ara	yen vwc	Dro	AGC	AGC	CTG	ACC	ACG	5844
	CCG (	CTG	CAC	AAC	GCA	ACC	GTG	CTC	GTC	GGT	AAG	~ m	~-~	CAC	ma m	C/D/C	
		Leu :	His	Asn	Ala	Thr	Val	Leu	Val	Glv	Lvs	Agn	Glu	CAG	TAI	CTG	5892
45	TTT (	CTG	GTC	CGG	GTG	GTG	ATG	CCC	TAC	CAA	CCC	000	TCC	TCG	GAC		5040
	Phe I	eu '	Val	Arg	• 44	AGT	Met	Pro	Tyr	Gln	Gly	Pro	Ser	Ser	Asn	Tur	5940
	GTG G	71U (	JTG Unit	AAG	ATG	ATC	CCG	GAC	AGC .	AGG	CTT	CCT	CCC	CGG			5988
	Val V	aı '		-,-		He	Pro	Asp	Ser .	Arg	Leu	Pro	Pro	Arg	His	Leu	3300
50	CAT G	la	Jal	UAC	ACC Th-	GGC .	AAG	ACC	TCG (	GCC	GTC	ATC	AAG	TGG	GAG	TCG	6036
	His A		1940		THE	стА	Dys	THE	ser /	Ala	Val	Ile	Lys	Trp	Glu	Ser	
	CCC T	AC C	SAC	ፐርጥ	ССТ	CAC .	CNC	1945	CTC -	-	<b></b>		1950	1			
	CCC T	yr i	Asp	Ser	Pro	Acn	GI n	JAC Aco	CIG :	LTC	TAT	GCG	ATC	GCA	GTT	AAA	6084
	Pro T	•	-			٠.٠٠	- 11 ·	veb	nen 1	ne	ıyr	Ala	TTE	Ala	Val	Lys	

		1955	;				1960	)				1965					
	GAT	CTG	ATA	CGA	AAG	ACG	GAC	CGG	AGC	T-AC	ΛΑΑ	STC	AAG	TCC	CGC	AAC	6132
	Asp	Leu	Ile	Arg	Lys			Arg	Ser	Tyr	Lys	Val	Lys	Ser	Arg	Asn	
5	1970	)				1975	5				1980		~~-			1985	6100
·	AGC	ACC	GTG	GAG	TAC	ACC	CTG	AGC	AAG	CTG	GAG	CCC	GGA	GGG	AAA	TAC	6180
	Ser	Thr	Val	Glu			Leu	Ser	rys	1999		Pro	GIA	GTA	2000		
	a. a	ama	3.000	ama	1990	) (TTC	ccc	AAC	ATC:			CAT	GCC	λCT			6228
	CAC	GIC	ATT	GIG	CAG	LOU	GGG	Asn	Mat	Ser	LVS	Agn	Δla	Ser	Val	LVa	0220
	HIS	val	TIE	2005		Leu	GLY	ASII	2010	)	2,5	1.00	nau	2019	;	_,,	
10	እጥር	<b>ACC</b>	ACC	GTT	TCG	TTA	TCG	GCA			GCC	TTA	AAA			ACA	6276
	Tle	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	Ile	Ile	Thr	
			2020	)				2025	5				2030	)			
	GAA	AAT	GAC	CAC	GTC	CTT	CTC	TTC	TGG	AAA	AGT	CTA	GCT	CTA	AAG	GAA	6324
	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys	Glu	
		2039	5				204	0				2045	5				6270
15	AAG	TAT	TTT	AAC	GAA	AGC	AGG	GGC	TAC	GAG	ATA	CAC	ATG	L.L.L.	GAT	AGC	6372
			Phe	Asn	Glu			Gly	Tyr	GIU	11e	HIS	Mec	Pne	Asp	2065	
	2050	·				205		CTT	CCC	א א ידי	2060		GAC	דממ	ጥጥር		6420
	GCC	ATG	AAT	ATC	ACC	GCA	TAC	Leu	Clv	Van	Thr	Thr	Asp	Agn	Phe	Phe	0420
	Ala	Met	Asn	TIE	207		ıyı	neu	Gry	207	5	****	wp	*****	208	0	
••	***	איזייני	TCC	AAC	CTG	DAG	ATG	GGT	CAC			ACA	TTC	ACG		CAG	6468
20	Tuc	TIA	Cor	Agn	Leu	Lvs	Met	Glv	His	Asn	Tyr	Thr	Phe	Thr	Val	Gln	
	-			2081	5				209	0				209	5		
	GCA	CGA	TGC	بلبلت	TTG	GGC	AGC	CAG	ATC	TGC	GGG	GAG	CCT	GCC	GTG	CTA	6516
	Ala	Arg	Cys	Leu	Leu	Gly	Ser	Gln	Ile	Cys	Gly	Glu	Pro	Ala	Val	Leu	
		_	210	n				210	5				211	0			c= c 4
25	CTG	TAT	GAT	GAG	CTG	GGG	TCT	GGT	GGC	GAT	GCG	TCG	GCG	ATG	CAG	GCT	6564
20	Leu			Glu	Leu	Gly	Ser	GIA	GTĀ	Asp	Ala	212	E	Met	GIII	Ala	
		211	5		~ · ·	ama	212	0	CITC	CTC	CTC			CTC	יואנייני	CTG	6612
	GCC	AGG	TCT	ACT	GAT	V-1	71-	772	Val	Val	Val	Pro	Tle	Leu	Phè	Leu	
	213		ser	1111	АЗР	213	5	AIG	Val	V 11 1	214	0				2145	
	213	U CTG	CTG	AGC	CTG	GGG	GTC	GGG	TTT	GCC	ATC	CTG	TAC	ACG	AAG	CAT	6660
30	Tle	Len	Leu	Ser	Leu	Glv	Val	Glv	Phe	Ala	Ile	Leu	Tyr	Thr	Lys	His	
					215	0				215	5				216	0	
	CGG	AGG	CTG	CAG	AGC	AGC	TTC	ACC	GCC	TTC	GCC	AAC	AGC	CAC	TAC	AGC	6708
	Arq	Arg	Leu	Gln	Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	His	TY	Ser	
	•	_		216	5				217	0				217	5		C756
	TCC	AGA	CTC	GGC	TCC	GCC	ATC	TTC	TCC	TCI	, GGG	GAT	GAC	TTG	GGG	GAG	6756
3 <b>5</b>	Ser	Arg			Ser	Ala	Ile	Phe	ser	ser	GIY	Asp	219	n Ten	GIY	/ Glu	
			218	0				218	5 D C7		THEFT	י ייירכ			· GTC	CCC	6804
	GAT	GAT	GAA	GAT	GCI	CCI	Mot	AIC TIC	The	GGA Gla	Dhe	Ser	Agr	Ast	Val	l Pro	
	Asp	_		Asp	ALa	Pro	220		. 1111	GIY	FIIC	220	5	, ,,,,,,			
	አጥር	219	כי גידיג י	GCC	TGZ	AAGA	GCT	TTCC	TCAC	TA G	AAAC			TGT	AAT	A	6856
40			. Ile			411.01	1001					-					
40	221	0															
	יידיד	רדבי	TGA	TAAA	GATA	GT 7	CGATO	GTT1	ra Ti	ATT	<b>LAAG</b>	TGC	ACT	TGA	GTT	GCAATA?	6916
	GTI	TTTA	TTA	TATO	GGC	CAA A	<b>VAAC</b>	AAAA	C A	LAAAA	AAAA	AAA A	AA				6961

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 300 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA to mRNA

55

50

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
_	ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCTGACC ATGGAGAGTG	60
5	ATATCAAGGT GAATGGCT.AT GTGGTGAACC TTTTCTGGGC ATTTGACACC CACAAGCAAG	120
	AGAGGAGAAC TTTGAACTTC CGAGGAAGCA TATTGTCACA CAAAGTTGGC AATCTGACAG	180
	CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGGGAT AGCCCTCTGG	240
10	CATTTGAGCA TGTTATGACC AGAGGGGTTC GCCCACCTGC ACCTAGCCTC AAGGCCAAAG	300
		300
	(2) INFORMATION FOR SEQ ID NO: 5:	
<b>15</b>	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 6642 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: double</li> <li>(D) TOPOLOGY: linear</li> </ul>	
	(ii) MOLECULE TYPE: cDNA to mRNA	
20		
	(vi) CEOUTING PROGRAM	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:	
25	ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCGT TCCTATTCAC CCTGGTCGCA	60
	CTGCTGCCGC CCGGAGCTCT CTGCGAAGTC TGGACGCAGA GGCTGCACGG CGGCAGCGCG	120
	CCCTTGCCCC AGGACCGGGG CTTCCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG	180
30	TGGGCGCGCG GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CCGGAGGAAA	240
	CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT	300
	TCCCACAATC AGATGGTGGT GCACTGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG	360
35	GCCCGAGATA GCCTGGCATT GGCGAGGCCC AAGAGCAGTG ATGTGTACGT GTCTTACGAC	420
	TATGGAAAAT CATTCAAGAA AATTTCAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT	480
	GAAGCTGTTA TCGCCCAGTT CTACCACAGC CCTGCGGACA ACAAGCGGTA CATCTTTGCA	540
40	GACGCTTATG CCCAGTACCT CTGGATCACG TTTGACTTCT GCAACACTCT TCAAGGCTTT	600
	TCCATCCCAT TTCGGGCAGC TGATCTCCTC CTACACAGTA AGGCCTCCAA CCTTCTCTTG	660
	GGCTTTGACA GGTCCCACCC CAACAAGCAG CTGTGGAAGT CAGATGACTT TGGCCAGACC	720
45	TGGATCATGA TTCAGGAACA TGTCAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA	780
	CCAAATACCA TCTACATTGA ACGACACGAA CCCTCTGGCT ACTCCACTGT CTTCCGAAGT	840
	ACAGATTTCT TCCAGTCCCG GGAAAACCAG GAAGTGATCC TTGAGGAAGT GAGAGATTTT	900
<i>E</i> 0	CAGCTTCGGG ACAAGTACAT GTTTGCTACA AAGGTGGTGC ATCTCTTGGG CAGTGAACAG	960
50	CAGTCTTCTG TCCAGCTCTG GGTCTCCTTT GGCCGGAAGC CCATGAGAGC AGCCCAGTTT	1020
	GTCACAAGAC ATCCTATTAA TGAATATTAC ATCGCAGATG CCTCCGAGGA CCAGGTGTTT	1080

	GTGTGTGTCA	GCCACAGTAA	CAACCGCACC	AATTTATACA	TCTCAGAGGC	<b>AGA(†GG(†CTG</b>	1140
	AAGTTCTCCC	TGTCCTTGGA	GAACGTGCTC	TATTACAGCC	CAGGAGGGGC	CGGCAGTGAC	1200
5 .	ACCTTGGTGA	GGTATTTTGC	AAATGAACCA	TTTGCTGACT	TCCACCGAGT	GGAAGGATTG	1260
	CAAGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCG	1320
	GTCATCACCT	TTGACAAAGG	GGGAACCTGG	GAGTTTCTTC	AGGCTCCAGC	CTTCACGGGA	1380
10	TATGGAGAGA	AAATCAATTG	TGAGCTTTCC	CAGGGCTGTT	CCCTTCATCT	GGCTCAGCGC	1440
	CTCAGTCAGC	TCCTCAACCT	CCAGCTCCGG	AGAATGCCCA	TCCTGTCCAA	GGAGTCGGCT	1500
	CCAGGCCTCA	TCATCGCCAC	TGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
15	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	TCACTACTAC	1620
	ACATGGGGAG	ACCACGGCGG	AATCATCACG	GCCATTGCCC	AGGGCATGGA	AACCAACGAG	1680
	CTAAAATACA	GTACCAATGA	AGGGGAGACC	TGGAAAACAT	TCATCTTCTC	TGAGAAGCCA	1740
20	GTGTTTGTGT	ATGGCCTCCT	CACAGAACCT	GGGGAGAAGA	GCACTGTCTT	CACCATCTTT	1800
	GGCTCGAACA	AAGAGAATGT	CCACAGCTGG	CTGATCCTCC	AGGTCAATGC	CACGGATGCC	1860
	TTGGGAGTTC	CCTGCACAGA	GAATGACTAC	AAGCTGTGGT	CACCATCTGA	TGAGCGGGGG	1920
25	AATGAGTGTT	TGCTGGGACA	CAAGACTGTT	TTCAAACGGC	GGACCCCCCA	TGCCACATGC	1980
	TTCAATGGAG	AGGACTTTGA	CAGGCCGGTG	GTCGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
	GACTATGAGT	GTGACTTCGG	TTTCAAGATG	AGTGAAGATT	TGTCATTAGA	GGTTTGTGTT	2100
30	CCAGATCCGG	AATTTTCTGG	AAAGTCATAC	TCCCCTCCTG	TGCCTTGCCC	TGTGGGTTCT	2160
	ACTTACAGGA	GAACGAGAGG	CTACCGGAAG	ATTTCTGGGG	ACACTTGTAG	CGGAGGAGAT	2220
•	GTTGAAGCGC	GACTGGAAGG	AGAGCTGGTC	CCCTGTCCCC	TGGCAGAAGA	GAACGAGTTC	2280
25	ATTCTGTATC	CTGTGAGGAA	ATCCATCTAC	CGCTATGACC	TGGCCTCGGG	AGCCACCGAG	2340
35	CAGTTGCCTC	TCACCGGGCT	ACGGGCAGCA	GTGGCCCTGG	ACTTTGACTA	TGAGCACAAC	2400
	TGTTTGTAT	r GGTCCGACCT	GGCCTTGGAC	GTCATCCAGC	GCCTCTGTTT	GAATGGAAGC	2460
	ACAGGGCAAG	G AGGTGATCAT	CAATTCTGGC	CTGGAGACAG	TAGAAGCTTT	GGCTTTTGAA	2520
40	CCCCTCAGC	C AGCTGCTTTA	CTGGGTAGAI	GCAGGCTTCA	AAAAGATTG/	A GGTAGCTAAT	258
	CCAGATGGC	G ACTTCCGACT	CACAATCGT	AATTCCTCTG	TGCTTGATC	TCCCAGGGCT	264
	CTGGTCCTC	g TGCCCCAAGA	GGGGGTGAT	TTCTGGACAG	ACTGGGGAG	A CCTGAAGCCT	270
45	GGGATTTAT	C GGAGCAATA	GGATGGTTC	C GCTGCCTATC	ACCTGGTGT	TGAGGATGTG	276
	AAGTGGCCC	A ATGGCATCT	TGTGGACGAC	CAGTGGATTI	ACTGGACGG	A TGCCTACCTG	282
	GAGTGCATA	G AGCGGATCA	GTTCAGTGG	C CAGCAGCGCT	CTGTCATTC	r ggacaacctc	288
50	CCGCACCCC	T ATGCCATTG	C TGTCTTTAA	G AATGAAATC	r actgggatg	A CTGGTCACAG	294
	CTCAGCATA	T TCCGAGCTT	C CAAATACAG	r gggtcccag	A TGGAGATTC	T GGCAAACCAG	300

	CTCACGGGGC TCATGGACAT GAAGATTTTC TACAAGGGGA AGAACACTGG AAGCAATSCC	3060
5	TGTGTGCCCA GGCCATGCAG CCTGCTGTGC CTGCCCAAGG CCAACAACAG TAGAAGCTGC	3120
5	AGGTGTCCAG AGGATGTGTC CAGCAGTGTG CTTCCATCAG GGGACCTGAT GTGTGACTGC	3180
	CCTCAGGGCT ATCAGCTCAA GAACAATACC TGTGTCAAAG AAGAGAACAC CTGTCTTCGC	3240
	AACCAGTATC GCTGCAGCAA CGGGAACTGT ATCAACAGCA TTTGGTGGTG TGACTTTGAC	3300
10	AACGACTGTG GAGACATGAG CGATGAGAGA AACTGCCCTA CCACCATCTG TGACCTGGAC	3360
	ACCCAGTTTC GTTGCCAGGA GTCTGGGACT TGTATCCCAC TGTCCTATAA ATGTGACCTT	3420
	GAGGATGACT GTGGAGACAA CAGTGATGAA AGTCATTGTG AAATGCACCA GTGCCGGAGT	3480
15	GACGAGTACA ACTGCAGTTC CGGCATGTGC ATCCGCTCCT CCTGGGTATG TGACGGGGAC	3540
	AACGACTGCA GGGACTGGTC TGATGAAGCC AACTGTACCG CCATCTATCA CACCTGTGAG	3600
	GCCTCCAACT TCCAGTGCCG AAACGGGCAC TGCATCCCCC AGCGGTGGGC GTGTGACGGG	3660
20	GATACGGACT GCCAGGATGG TTCCGATGAG GATCCAGTCA ACTGTGAGAA GAAGTGCAAT	3720
	GGATTCCGCT GCCCAAACGG CACTTGCATC CCATCCAGCA AACATTGTGA TGGTCTGCGT	3780
	GATTGCTCTG ATGGCTCCGA TGAACAGCAC TGCGAGCCCC TCTGTACGCA CTTCATGGAC	3840
25	TTTGTGTGTA AGAACCGCCA GCAGTGCCTG TTCCACTCCA TGGTCTGTGA CGGAATCATC	3900
	CAGTGCCGCG ACGGGTCCGA TGAGGATGCG GCGTTTGCAG GATGCTCCCA AGATCCTGAG	3960
	TTCCACAAGG TATGTGATGA GTTCGGTTTC CAGTGTCAGA ATGGAGTGTG CATCAGTTTG	4020
30	ATTTGGAAGT GCGACGGGAT GGATGATTGC GGCGATTATT CTGATGAAGC CAACTGCGAA	4080
-	AACCCCACAG AAGCCCCAAA CTGCTCCCGC TACTTCCAGT TTCGGTGTGA GAATGGCCAC	4140
	TGCATCCCCA ACAGATGGAA ATGTGACAGG GAGAACGACT GTGGGGACTG GTCTGATGAG	4200
35	AAGGATTGTG GAGATTCACA TATTCTTCCC TTCTCGACTC CTGGGCCCTC CACGTGTCTG	4260
	CCCAATTACT ACCGCTGCAG CAGTGGGACC TGCGTGATGG ACACCTGGGT GTGCGACGGG	4320
	TACCGAGATT GTGCAGATGG CTCTGACGAG GAAGCCTGCC CCTTGCTTGC AAACGTCACT	4380
40	GCTGCCTCCA CTCCCACCCA ACTTGGGCGA TGTGACCGAT TTGAGTTCGA ATGCCACCAA	4440
	CCGAAGACGT GTATTCCCAA CTGGAAGCGC TGTGACGGCC ACCAAGATTG CCAGGATGGC	4500
	CGGGACGAGG CCAATTGCCC CACACACAGC ACCTTGACTT GCATGAGCAG GGAGTTCCAG	4560
<b>4</b> 5	TGCGAGGACG GGGAGGCCTG CATTGTGCTC TCGGAGCGCT GCGACGGCTT CCTGGACTGC	4620
40	TCGGACGAGA GCGATGAAAA GGCCTGCAGT GATGAGTTGA CTGTGTACAA AGTACAGAAT	4680
	CTTCAGTGGA CAGCTGACTT CTCTGGGGAT GTGACTTTGA CCTGGATGAG GCCCAAAAAA	4740
	ATGCCCTCTG CATCTTGTGT ATATAATGTC TACTACAGGG TGGTTGGAGA GAGCATATGG	4800
50	AAGACTCTGG AGACCCACAG CAATAAGACA AACACTGTAT TAAAAGTCTT GAAACCAGAT	4860
	ACCACGTATC AGGTTAAAGT ACAGGTTCAG TGTCTCAGCA AGGCACACAA CACCAATGAC	4920

	TTTGTGACCC	TGAGGACCCC	AGAGGGATTG	CCAGATGCCC	CTCGAAATCT	CCAGCTGTCA	4980
	CTCCCCAGGG	AAGCAGAAGG	TGTGATTGTA	GGCCACTGGG	CTCCTCCCAT	CCACACCCAT	5040
5	GGCCTCIATCC	GTGAGTACAT	TGTAGAATAC	AGCAGGAGTG	GTTCCAAGAT	GTGGGCCTCC	5100
	CAGAGGGCTG	CTAGTAACTT	TACAGAAATC	AAGAACTTAT	TGGTCAACAC	TCTATACACC	5160
	GTCAGAGTGG	CTGCGGTGAC	TAGTCGTGGA	ATAGGAAACT	GGAGCGATTC	TAAATCCATT	5220
10	ACCACCATAA	AAGGAAAAGT	GATCCCACCA	CCAGATATCC	ACATTGACAG	CTATGGTGAA	5280
	AATTATCTAA	GCTTCACCCT	GACCATGGAG	AGTGATATCA	AGGTGAATGG	CTATGTGGTG	5340
	AACCTTTTCT	GGGCATTTGA	CACCCACAAG	CAAGAGAGGA	GAACTTTGAA	CTTCCGAGGA	5400
15	AGCATATTGT	CACACAAAGT	TGGCAATCTG	ACAGCTCATA	CATCCTATGA	GATTTCTGCC	5460
	TGGGCCAAGA	CTGACTTGGG	GGATAGCCCT	CTGGCATTTG	AGCATGTTAT	GACCAGAGGG	5520
	GTTCGCCCAC	CTGCACCTAG	CCTCAAGGCC	AAAGCCATCA	ACCAGACTGC	AGTGGAATGT	5580
20	ACCTGGACCG	GCCCCGGAA	TGTGGTTTAT	GGTATTTTCT	ATGCCACGTC	CTTTCTTGAC	5640
	CTCTATCGCA	ACCCGAAGAG	CTTGACTACT	TCACTCCACA	ACAAGACGGT	CATTGTCAGT	5700
	AAGGATGAGC	AGTATTTGTT	TCTGGTCCGT	GTAGTGGTAC	CCTACCAGGG	GCCATCCTCT	5760
25	GACTACGTTG	TAGTGAAGAT	GATCCCGGAC	AGCAGGCTTC	CACCCCGTCA	CCTGCATGTG	5820
25	GTTCATACGG	GCAAAACCTC	CGTGGTCATC	AAGTGGGAAT	CACCGTATGA	CTCTCCTGAC	5880
	CAGGACTTGT	TGTATGCAAT	TGCAGTCAAA	GATCTCATAA	GAAAGACTGA	CAGGAGCTAC	5940
22	AAAGTAAAAT	CCCGTAACAG	CACTGTGGAA	TACACCCTTA	ACAAGTTGGA	GCCTGGCGGG	6000
30	AAATACCACA	TCATTGTCCA	ACTGGGGAAC	ATGAGCAAAG	ATTCCAGCAT	AAAAATTACC	6060
	ACAGTTTCAT	TATCAGCACC	TGATGCCTTA	AAAATCATAA	CAGAAAATGA	TCATGTTCTT	6120
	CTGTTTTGGA	AAAGCCTGGC	TTTAAAGGAA	AAGCATTTTA	ATGAAAGCAG	GGGCTATGAG	6180
35	ATACACATGT	TTGATAGTGC	CATGAATATC	ACAGCTTACC	TTGGGAATAC	TACTGACAAT	6240
	TTCTTTAAAA	TTTCCAACCT	GAAGATGGGT	CATAATTACA	CGTTCACCGT	CCAAGCAAGA	6300
	TGCCTTTTTG	GCAACCAGAT	CTGTGGGGAG	CCTGCCATCC	TGCTGTACGA	TGAGCTGGGG	6360
40	TCTGGTGCAG	ATGCATCTGC	AACGCAGGCT	GCCAGATCTA	CGGATGTTG	TGCTGTGGTG	6420
	GTGCCCATCT	TATTCCTGAT	ACTGCTGAGO	CTGGGGGTGG	GGTTTGCCAT	CCTGTACACG	6480
	AAGCACCGGA	GGCTGCAGAG	CAGCTTCACC	GCCTTCGCCA	ACAGCCACT	A CAGCTCCAGG	6540
45	CTGGGGTCCG	CAATCTTCTC	CTCTGGGGAT	GACCTGGGGG	AAGATGATG	A AGATGCCCCT	6600
	ATGATAACTG	GATTTTCAGA	TGACGTCCCC	: ATGGTGATAG	cc cc		6642

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2214 amino acids
(B) TYPE: amino acid

55

(C) STRANDEDNESS:
(D) TOPOLOGY: linear

5	(ii)	MOI	ECUI	LE TY	PE:	Prot	ein								•	
	(xi)	SEÇ	UENC	E DE	SCRI	PTIC	N: 5	SEQ I	D NO	): 6:						
10		Ala										J Leu	Pro	Phe	Leu 15	Phe
•	Thr	Leu	Val	Ala 20	Leu	Leu	Pro	Pro	Gly 25	Ala	Lev	. Cys	Glu	. Val	Trp	Thi
15	Gln	Arg	Leu 35	His	Gly	Gly	Ser	Ala 40	Pro	Leu	Pro	Gln	Asp 45	Arg	Gly	Phe
	Leu	Val 50	Val	Gln	Gly	Asp	Pro 55	Arg	Glu	Leu	Arg	Leu 60	Trp	Ala	Arg	Gly
20		Ala				70					75					80
		Ser			0.5					90					95	
25		Leu		100					105					110		
		Ser	~					120					125			•
30		Pro 130					135					140				
	-10	Lys				120					155					160
35		Ala			105					170					175	
		Ile		100					192					190		
40		Cys	195					200					205			
		Leu 210					215					220				
45		His				250					235					240
		Ile			443					250					255	
50		Tyr		200					265					270		
	GIY	Tyr	Ser 275	Thr	Val	Phe	Arg	Ser 280	Thr	Asp	Phe	Phe	Gln 285	Ser	Arg	Glu

		Asn	Gln 290	Glu	Val	Ile	Leu	Glu 295	Glu	Val	Arg	qzA	300 5µs	Cln	Leu	Arg	Asp
5		Lys 305	Tyr	Met	Phe	Ala	Thr 310	Lys	Val	Val	His	Leu 315	Leu	Gly	Ser	Glu	Gln 320
		Gln	Ser	Ser	Val	Gln 325	Leu	Trp	Val	Ser	Phe 330	Gly	Arg	Lys	Pro	Met 335	Arg
10	~	Ala	Ala	Gln	Phe 340	Val	Thr	Arg	His	Pro 345	Ile	Asn	Glu	Tyr	Tyr 350	Ile	Ala
		Asp	Ala	Ser 355	Glu	Asp	Gln	Val	Phe 360	Val	Сув	Val	Ser	His 365	Ser	Asn	Asn
15		Arg	Thr 370	Asn	Leu	Tyr	Ile	Ser 375	Glu	Ala	Glu	Gly	Leu 380	Lys	Phe	Ser	Leu
		Ser 385	Leu	Glu	Asn	Val	Leu 390	Tyr	Tyr	Ser	Pro	Gly 395	Gly	Ala	Gly	Ser	Asp 400
20		Thr	Leu	Val	Arg	Tyr 405	Phe	Ala	Asn	Glu	Pro 410	Phe	Ala	Asp	Phe	His 415	Arg
		Val	Glu	Gly	Leu 420	Gln	Gly	Val	Tyr	Ile 425	Ala	Thr	Leu	Ile	Asn 430	Gly	Ser
25		Met	Asn	Glu 435	Glu	Asn	Met	Arg	Ser 440	Val	Ile	Thr	Phe	Asp 445	Lys	Gly	Gly
		Thr	Trp 450	Glu	Phe	Leu	Gln	Ala 455	Pro	Ala	Phe	Thr	Gly 460	Tyr	Gly	Glu	Lys
3 <i>0</i>		Ile 465	Asn	Cys	Glu	Leu	Ser 470	Gln	Gly	Cys	Ser	Leu 475	His	Leu	Ala	Gln	Arg 480
		Leu	Ser	Gln	Leu	Leu 485	Asn	Leu	Gln	Leu	Arg 490	Arg	Met	Pro	Ile	Leu 495	Ser
3 <i>5</i>		Lys	Glu	Ser	Ala 500	Pro	Gly	Leu	Ile	Ile 505	Ala	Thr	Gly	Ser	Val 510	Gly	Lys
		Asn	Leu	Ala 515	Ser	Lys	Thr	Asn	Val 520	Tyr	Ile	Ser	Ser	Ser 525	Ala	Gly	Ala
40		Arg	Trp 530	Arg	Glu	Ala	Leu	Pro 535	Gly	Pro	His	Tyr	Tyr 540	Thr	Trp	Gly	Asp
40		His 5 <b>4</b> 5	Gly	Gly	Ile	Ile	Thr 550	Ala	Ile	Ala	Gln	Gly 555	Met	Glu	Thr	Asn	Glu 560
		Leu	Lys	Tyr	Ser	Thr 565	Asn	Glu	Gly	Glu	Thr 570	Trp	Lys	Thr	Phe	Ile 575	Phe
45		Ser	Glu	Lys	Pro 580	Val	Phe	Val	Tyr	Gly 585	Leu	Leu	Thr	Glu	Pro 590	Gly	Glu
		Lys	Ser	Thr 595	Val	Phe	Thr	Ile	Phe 600	Gly	Ser	Asn	Lys	Glu 605	Asn	Val	His
50		Ser	Trp 610	Leu	Ile	Leu	Gln	Val 615	Asn	Ala	Thr	Asp	Ala 620	Leu	Gly	Val	Pro
		Cys	Thr	Glu	Asn	Asp	Tyr	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly

	625			630					635					640
5	Asn Gl	ı Cys Le	Leu 645	Gly	His	Lys	Thr	Val 650	Phe	Lys	Arg	_	Thr 655	Pro
	His Ala	a Thr Cy:		Asn	Gly	Glu	Asp 665	Phe	Asp	Arg	Pro	Val 670	Val	Val
10	Ser Ası	n Cys Se: 675	Cys	Thr	Arg	Glu 680	Asp	Tyr	Glu	Cys	Asp 685	Phe	Gly	Phe
	Lys Met	Ser Gl	Asp	Leu	Ser 695	Leu	Glu	Val	Cys	Val 700	Pro	Asp	Pro	Glu
15	Phe Set 705	c Gly Ly	Ser	Tyr 710	Ser	Pro	Pro	Val	Pro 715	Cys	Pro	Val	Gly	Ser 720
	Thr Ty:	r Arg Ar	725	Arg	Gly	Tyr	Arg	Lys 730	Ile	Ser	Gly	Asp	Thr 735	Суз
00	Ser Gl	y Gly Asy 74		Glu	Ala	Arg	Leu 745	Glu	Gly	Glu	Leu	Val 750	Pro	Суѕ
20	Pro Le	ı Ala Glı 755	Glu	Asn	Glu	Phe 760	Ile	Leu	Tyr	Ala	Val 765	Arg	Lys	Ser
	Ile Ty	r Arg Ty: )	Asp	Leu	Ala 775	Ser	Gly	Ala	Thr	Glu 780	Gln	Leu	Pro	Leu
25	Thr Gly	/ Leu Ar	, Ala	<b>Ala</b> 790	Val	Ala	Leu	Asp	Phe 795	Asp	Tyr	Glu	His	Asn 800
	Cys Le	ı Tyr Tr	Ser 805	Asp	Leu	Ala	Leu	Asp 810	Val	Ile	Gln	Arg	Leu 815	Cys
30	Leu Ası	n Gly Se 82		Gly	Gln	Glu	Val 825	Ile	Ile	Asn	Ser	Gly 830	Leu	Glu
	Thr Va	l Glu Al 835	ı Leu	Ala	Phe	Glu 840	Pro	Leu	Ser	Gln	Leu 845	Leu	Tyr	Trp
35	Val Ası 85	o Ala Gl	y Phe	Lys	Lys 855	Ile	Glu	Val	Ala	Asn 860	Pro	Asp	Gly	Asp
	Phe Are	g Leu Th		Val 870	Asn	Ser	Ser	Val	Leu 875	Asp	Arg	Pro	Arg	Ala 880
40	Leu Va	l Leu Va	885	Gln	Glu	Gly	Val	Met 890	Phe	Trp	Thr	Asp	Trp 895	Gly
	Asp Le	u Lys Pr 90		Ile	Tyr	Arg	Ser 905	Asn	Met	Asp	Gly	Ser 910	Ala	Ala
45	Tyr Hi	s Leu Va 915	l Ser	Glu	Asp	Val 920	Lys	Trp	Pro	Asn	Gly 925	Ile	Ser	Val
	Aap As 93	p Gln Tr O	) Ile	Tyr	Trp 935	Thr	Asp	Ala	Tyr	Leu 940	Glu	СЛв	Ile	Glu
50	Arg Il	e Thr Ph	e Ser	Gly 950	Gln	Gln	Arg	Ser	Val 955	Ile	Leu	Asp	Asn	Leu 960
	Pro Hi	s Pro Ty	965		Ala	Val	Phe	Lys 970	Asn	Glu	Ile	Tyr	Trp 975	Asp

	Asp	Trp	Ser	Gln 980	Leu	Ser	Ile	Phe	Arg 985	Ala	Sex	Lуя	Tyr	Ser 990	Gly	Ser
5	Gln	Met	Glu 995	Ile	Leu	Ala	Asn	Gln 1000		Thr	Gly	Leu	Met 1005		Met	Lys
	Ile	Phe 1010		Lys	Gly	Lys	Asn 1015		Gly	Ser	Asn	Ala 1020		Val	Pro	Arg
10	Pro 1025		Ser	Leu	Leu	Cys 1030		Pro	Lys	Ala	Asn 1035		Ser	Arg	Ser	Cys 1040
	Arg	Cys	Pro	Glu	Asp 1045		Ser	Ser	Ser	Val 1050		Pro	Ser	Gly	Asp 1055	
15	Met	Cys	Asp	Cys 1060		Gln	Gly	Tyr	Gln 1065		Lys	Asn	Asn	Thr 1070		Val
	Lys	Glu	Glu 1075		Thr	Cys	Leu	Arg 1080		Gln	Tyr	Arg	Cys 1085		Asn	Gly
20	Asn	Cys 1090		Asn	Ser	Ile	Trp 1095		Cys	Asp	Phe	Asp 1100		Asp	Cys	Gly
	Asp 1105		Ser	Asp	Glu	Arg 1110		Сув	Pro	Thr	Thr 1115		Cys	Asp	Leu	Asp 1120
25	Thr	Gln	Phe	Arg	Cys 1125		Glu	Ser	Gly	Thr 1130		Ile	Pro	Leu	Ser 1135	
	Lys	Cys	Asp	Leu 1140		Asp	Asp	Cys	Gly 1145		Asn	Ser	Asp	Glu 1150		His
30	Cys	Glu	Met 1155		Gln	Cys	Arg	Ser 1160		Glu	Tyr	Asn	Cys 1165		Ser	Gly
	Met	Cys 1170		Arg	Ser	Ser	Trp 1175		Cys	Asp	Gly	Asp 1180		Asp	Cys	Arg
35	Asp 1185		Ser	Asp	Glu	Ala 1190		Cys	Thr	Ala	Ile 1195		His	Thr	Сув	Glu 1200
	Ala	Ser	Asn	Phe	Gln 1205		Arg	Asn	Gly	His 1210		Ile	Pro	Gln	Arg 1215	
40	Ala	Cys	Asp	Gly 1220	_	Thr	Asp	Cys	Gln 1225	_	Gly	Ser	Asp	Glu 1230	Asp	Pro
••	Val	Asn	Cys 1235		Lys	Lys	Cys	Asn 1240		Phe	Arg	Суѕ	Pro 1245		Gly	Thr
	Cys	Ile 1250		Ser	Ser	Lys	His 1255		Asp	Gly	Leu	Arg 1260	_	Суз	Ser	Asp
<b>1</b> 5	Gly 1265		Asp	Glu	Gln	His 1270		Glu	Pro	Leu	Cys 1275		His	Phe	Met	Asp 1280
	Phe	Val	Cys	Lys	Asn 1285		Gln	Gln	Суз	Leu 1290		His	Ser	Met	Val 1295	
50	Asp	Gly	Ile	Ile 1300		Cys	Arg	Asp	Gly 1305		Asp	Glu	Asp	Ala 1310	Ala	Phe

	Ala	Gly	Cys 1319	Ser 5	Gln	Asp	Pro	Glu 1320	Phe )	His	Lys	Ľs)	Сув 1329		Clu	Phe
5	Gly	Phe 1330	Gln )	Cys	Gln	Asn	Gly 1335	Val	Cys	Ile	Ser	Leu 1340		Trp	Lys	Cys
	Asp 1345	Gly 5	Met	Asp	Asp	Cys 1350	Gly )	Asp	Tyr	Ser	Asp 1355		Ala	Asn	Cys	Glu 1360
10	Asn	Pro	Thr	Glu	Ala 1369	Pro	Asn	Cys	Ser	Arg 1370		Phe	Gln	Phe	Arg 1379	
	Glu	Asn	Gly	His 1380	Cys )	Ile	Pro	Asn	Arg 1389	Trp	Lys	Суз	Asp	Arg 1390		Asn
15	Asp	Суз	Gly 1395	Asp	Trp	Ser	Asp	Glu 1400	Lys )	Asp	Cys	Gly	Asp 1405		His	Ile
	Leu	Pro 1410	Phe	Ser	Thr	Pro	Gly 1415		Ser	Thr	Cys	Leu 1420		Asn	Tyr	Tyr
20	Arg 1425	Cys	Ser	Ser	Gly	Thr 1430	Cys )	Val	Met	Asp	Thr 1439		Val	Cys	Asp	Gly 1440
	Tyr	Arg	Asp	Cys	Ala 1445		Gly	Ser	Asp	Glu 1450		Ala	Cys	Pro	Leu 1455	
25	Ala	Asn	Val	Thr 1460	Ala	Ala	Ser	Thr	Pro 1465		Gln	Leu		Arg 1470		Asp
	Arg	Phe	Glu 1475		Glu	Сув	His	Gln 1480	Pro	Lys	Thr	Суз	Ile 1485		Asn	Trp
30	Lys	Arg 1490	Cys )	Asp	Gly	His	Gln 1495		Cys	Gln	Asp	Gly 1500		Asp	Glu	Ala
	Asn 1505	Cys	Pro	Thr	His	Ser 1510	Thr	Leu	Thr	Суз	Met 1515		Arg	Glu	Phe	Gln 1520
35	Cys	Glu	Asp	Gly	Glu 1525		Сув	Ile	Val	Leu 1530		Glu	Arg	Сув	Asp 1535	_
	Phe	Leu	Asp	Cys 1540	Ser	Asp	Glu	Ser	Asp 1545		Lys	Ala	Cys	Ser 1550		Glu
40	Leu	Thr	Val 1555	Tyr	Lys	Val	Gln	Asn 1560		Gln	Trp	Thr	Ala 1565		Phe	Ser
	Gly	Asp 1570	Val	Thr	Leu	Thr	Trp 1575		Arg	Pro	Lys	Lys 1580		Pro	Ser	Ala
	Ser 1585		Val	Tyr	Asn	Val 1590		Tyr	Arg	Val	Val 1595		Glu	Ser	Ile	Trp 1600
45	Lys	Thr	Leu	Glu	Thr 1605		Ser	Asn	Lys	Thr 1610		Thr	Val	Leu	Lys 1619	
	Leu	Lys	Pro	<b>Asp</b> 1620		Thr	Tyr	Gln	Val 1629		Val	Gln	Val	Gln 1630		Leu
50	Ser	Lys	Ala 1635		Asn	Thr	Asn	Asp 1640		Val	Thr	Leu	Arg 1645		Pro	Glu
	Gly	Leu	Pro	Asp	Ala	Pro	Arg	Asn	Leu	Gln	Leu	Ser	Leu	Pro	Arg	Glu

		1650	)				1655	;				3.650	)			
5	Ala 1665		Gly	Val	Ile	Val 1670		His	Trp	Ala	Pro 1675		Ile	His	Thr	His 1680
	Gly	Leu	Ile	Arg	Glu 1685		Ile	Val	Glu	Tyr 1690		Arg	Ser	Gly	Ser 1695	
10	Met	Trp	Ala	Ser 1700		Arg	Ala	Ala	Ser 1705		Phe	Thr	Glu	Ile 1710	-	Asn
	Leu	Leu	Val 1715		Thr	Leu	Tyr	Thr 1720		Arg	Val	Ala	Ala 1725		Thr	Ser
	Arg	Gly 1730		Gly	Asn	Trp	Ser 1735	-	Ser	Lys	Ser	Ile 1740		Thr	Ile	Lys
15	Gly 1745	-	Val	Ile	Pro	Pro 1750		Asp	Ile	His	Ile 1755		Ser	Tyr	Gly	Glu 1760
	Asn	Tyr	Leu	Ser	Phe 1765		Leu	Thr	Met	Glu 1770		Asp	Ile	Lys	Val 1775	
20	Gly	Tyr	Val	Val 1780		Leu	Phe	Trp	Ala 1785		Asp	Thr	His	Lys 1790		Glu
	Arg	Arg	Thr 1795		Asn	Phe	Arg	Gly 1800		Ile	Leu	Ser	His 1805		Val	Gly
25	Asn	Leu 1810		Ala	His	Thr	Ser 1815		Glu	Ile	Ser	Ala 1820		Ala	Lys	Thr
	<b>Asp</b> 1825		Gly	Asp	Ser	Pro 1830		Ala	Phe	Glu	His 1835		Met	Thr	Arg	Gly 1840
30	Val	Arg	Pro	Pro	Ala 1845		Ser	Leu	Lys	Ala 1850		Ala	Ile	Asn	Gln 1855	
	Ala	Val	Glu	Cys 1860	Thr	Trp	Thr	Gly	Pro 1865		Asn	Val	Val	Tyr 1870		Ile
35	Phe	Tyr	Ala 1879		Ser	Phe	Leu	Asp 1880		Tyr	Arg	Asn	Pro 1885		Ser	Leu
	Thr	Thr 1890		Leu	His	Asn	Lys 1899		Val	Ile	Val	Ser 1900		Asp	Glu	Gln
40	Tyr 1905		Phe	Leu	Val	Arg 1910		Val	Val	Pro	Tyr 191		Gly	Pro	Ser	Ser 1920
	Asp	Tyr	Val	Val	Val 1929		Met	Ile	Pro	Asp 193		Arg	Leu	Pro	Pro 1935	
45	His	Leu	His	Val 194	Val	His	Thr	Gly	Lys 194		Ser	Val	Val	Ile 1950		Trp
	Glu	Ser	Pro 195		Asp	Ser	Pro	Asp 1960		Asp	Leu	Leu	Tyr 196		Ile	Ala
50	Val	Lys 1970		Leu	Ile	Arg	Lys 197		Asp	Arg	Ser	Tyr 198		Val	Lys	Ser
	Arg 1985		Ser	Thr	Val	Glu 199		Thr	Leu	Asn	Lys 199		Glu	Pro	Gly	Gly 2000

		Lys	Tyr	His	Ile	Ile 200	Val	Gln	Leu	Gly	Asn 201	Met O	Ser	Lys	Asp	Ser 2019	
5		Ile	Lys	Ile	Thr 202	Thr 0	Val	Ser	Leu	Ser 202	Ala 5	Pro	Asp	Ala	Leu 2030		Ile
		Ile	Thr	Glu 203	Asn 5	qaA	His	Val	Leu 2040	Leu	Phe	Trp	Lys	Ser 2045	Leu 5	Ala	Leu
10		Lys	Glu 2050	Lys 0	His	Phe	Asn	Glu 205	Ser 5	Arg	Gly	Tyr	Glu 2060		His	Met	Phe
		Asp 2065	Ser 5	Ala	Met	Asn	Ile 2070	Thr	Ala	Tyr	Leu	Gly 2075	Asn 5	Thr	Thr	Asp	Asn 2080
15		Phe	Phe	Lys	Ile	Ser 208	Asn	Leu	Lys	Met	Gly 2090		Asn	Tyr	Thr	Phe 2095	
		Val	Gln	Ala	Arg 210	Cys	Leu	Phe	Gly	Asn 2105		Ile	Cys	Gly	Glu 2110		Ala
20		Ile	Leu	Leu 2119	Tyr	Asp	Glu	Leu	Gly 2120	Ser	Gly	Ala	Asp	Ala 2125	Ser	Ala	Thr
		Gln	Ala 2130	Ala	Arg	Ser	Thr	Asp 2135		Ala	Ala	Val	Val 2140		Pro	Ile	Leu
25		Phe 2145	Leu	Ile	Leu	Leu	Ser 2150	Leu )	Gly	Val	Gly	Phe 2155		Ile	.Leu	Tyr	Thr 2160
		Lys	His	Arg	Arg	Leu 2165	Gln	Ser	Ser	Phe	Thr 2170		Phe	Ala	Asn	Ser 2175	
30		Tyr	Ser	Ser	Arg 2180	Leu )	Gly	Ser	Ala	Ile 2185		Ser	Ser	Gly	Asp 2190		Leu
-		Gly	Glu	Asp 2195	Asp	Glu	Asp	Ala	Pro 2200		Ile	Thr	Gly	Phe 2205	Ser	Asp	Asp
35		Val	Pro 2210	Met )	Val	Ile	Ala		•								
	(2)	INFOR	ITAM	ON F	OR S	EQ I	D NC	): 7:									
40		(i)	(A) (B) (C)	IENCE LEN TYF STR TOF	IGTH: PE: I PANDE	684 ucle	3 ba ic a SS: d	se p cid loubl	airs	3							
		(ii)	MOLE	CULE	TYF	E: c	DNA	to m	RNA								
45		(ix)	(A)	NAM	IE/KE				.de								
50		(ix)	(C) FEAT (A) (B)	LOC IDE URE: NAM LOC IDE	NTIF	FICAT EY: π ON:16	ION at p	METH epti	de								
55																	

# (xi) SEQUENCE DESCRIPTION: SEQ ID MO: 7:

5	TCG	CGCT	GCA (	CATTO	CTCT	CC TO	GCG	GCGG	C GC	CACC:	CCG FGCA	GCC( GTA(	CAGCO	GC (	rctcg/	TGGCC AACATG Met	23 83
	Ala	Thr	Arg	Ser 5	AGC Ser	Arg	Arg	Glu	Ser 10	Arg	Leu	Pro	Phe	Leu 15	Phe	Thr	131
10	Leu	Val	Ala 20	Leu	CTG Leu	Pro	Pro	Gly 25	Ala	Leu	Cys	Glu	Val 30	Trp	Thr	Gln	179
	AGG Arg	CTG Leu 35	CAC His	GGC Gly	GGC Gly	AGC Ser	GCG Ala 40	CCC Pro	TTG Leu	CCC Pro	CAG Gln	GAC Asp 45	CGG Arg	GGC Gly	TTC Phe	CTC Leu	227
15	GTG Val 50	GTG Val	CAG Gln	GGC Gly	GAC Asp	CCG Pro 55	CGC Arg	GAG Glu	CTG Leu	CGG Arg	CTG Leu 60	TGG Trp	GCG Ala	CGC <b>Arg</b>	GGG Gly	GAT Asp 65	275
	GCC Ala	AGG Arg	GGG Gly	GCG Ala	AGC Ser 70	CGC Arg	GCG Ala	GAC Asp	GAG Glu	AAG Lys 75	CCG Pro	CTC Leu	CGG Arg	AGG Arg	AAA Lys 80	CGG	323
20	AGC Ser	GCT Ala	GCC Ala	CTG Leu 85	CAG Gln	CCC Pro	GAG Glu	CCC Pro	ATC Ile 90	AAG Lys	GTG Val	TAC Tyr	GGA Gly	CAG Gln 95	GTT Val	AGT Ser	371
	CTG Leu	AAT Asn	GAT Asp 100	TCC	CAC His	AAT Asn	CAG Gln	ATG Met 105	GTG	GTG Val	CAC His	TGG Trp	GCT Ala 110	GGA	GAG Glu	AAA Lys	419
25	AGC Ser	AAC Asn 115	GTG	ATC Ile	GTG Val	GCC Ala	TTG Leu 120	GCC	CGA Arg	GAT Asp	AGC Ser	CTG Leu 125	GCA	TTG Leu	GCG Ala	AGG Arg	467
	CCC Pro 130	AAG	AGC Ser	AGT Ser	GAT Asp	GTG Val 135	TAC	GTG Val	TCT Ser	TAC Tyr	GAC Asp 140	TAT	GGA Gly	AAA Lys	TCA Ser	TTC Phe 145	515
30	AAG	Lys AAA	ATT Ile	TCA Ser	GAC Asp 150	AAG	TTA Leu	AAC Asn	TTT Phe	GGC Gly 155	TTG	GGA Gly	AAT Asn	AGG Arg	AGT Ser 160	GAA	563
	GCT Ala	GTT Val	ATC Ile	GCC Ala 165	CAG Gln	TTC Phe	TAC Tyr	CAC His	AGC Ser 170	CCT	GCG Ala	GAC Asp	AAC Asn	AAG Lys 175	CGG	TAC Tyr	611
				GAC	GCT Ala				TAC				Thr	TTT			659
35	TGC Cys	Asn	ACT	CTT Leu	CAA Gln	GGC Gly	Phe	TCC	ATC Ile	CCA Pro	TTT Phe	Arg	190 GCA Ala	GCT Ala	GAT Asp	CTC Leu	70 <b>7</b>
	CTC Leu 210	195 CTA Leu	CAC His	AGT Ser	AAG Lys	Ala	200 TCC Ser	AAC Asn	CTT Leu	CTC Leu	Leu	GGC Gly	TTT Phe	GAC Asp	AGG Arg	Ser	755
40	CAC	CCC Pro	AAC Asn	AAG Lys	CAG Gln	215 CTG Leu	TGG Trp	AAG Lys	TCA Ser	Asp	GAC Asp	TTT Phe	GGC Gly	CAG Gln	Thr	225 TGG Trp	803
	ATC Ile	ATG Met	ATT Ile	Gln	GAA Glu	CAT His	GTC Val	AAG Lys	Ser	235 TTT Phe	TCT Ser	TGG Trp	GGA Gly	Ile	GAT Asp	CCC Pro	851
45			Lys		AAT Asn			Tyr					Glu				899
	TAC Tyr	Ser	260 ACT Thr	GTC Val	TTC Phe	CGA Arg	Ser	265 ACA Thr	GAT Asp	TTC Phe	TTC Phe	Gln	270 TCC Ser	CGG Arg	GAA Glu	AAC Asn	947
50	CAG Gln 290	275 GAA Glu	GTG Val	ATC Ile	CTT Leu	GAG Glu 295	280 GAA Glu	GTG Val	AGA Arg	GAT Asp	TTT Phe 300	285 CAG Gln	CTT Leu	CGG Arg	GAC Asp	AAG Lys 305	995

	TAC Tyr	ATG Met	TTT Phe	GCT Ala	ACA Thr 310	AAG Lys	GTG Val	GTG Val	CAT His	CTC Leu 315	TTG Leu	Gl?	ACT Ser	GAA Glu	CAC Gln 320	CAG Cln	1943
5	TCT Ser	TCT Ser	GTC Val	CAG Gln 325	CTC Leu	TGG Trp	GTC Val	TCC Ser	TTT Phe 330	GGC	CGG Arg	AAG Lys	CCC Pro	ATG Met 335	AGA	GCA Ala	1091
	GCC Ala	CAG Gln	TTT Phe 340	GTC	ACA Thr	AGA Arg	CAT His	CCT Pro 345	ATT	AAT Asn	GAA Glu	TAT Tyr	TAC Tyr 350	ATC	GCA Ala	GAT Asp	1139
10								GTG Val									1187
	Thr 370	Asn	Leu	Tyr	Ile	<b>Ser</b> 375	Glu	GCA Ala	Glu	Gly	Leu 380	Lys	Phe	Ser	Leu	Ser 385	1235
15								AGC Ser									1283
								GAA Glu									1331
20								ATT Ile 425									1379
								GTC Val									1427
<i>2</i> 5								GCC Ala									1475
								TGT Cys									1523
30								CTC Leu									1571
								ATC 1le 505									1619
35	Leu	Ala 515	Ser	Lys	Thr	Asn	Val 520	TAC Tyr	Ile	Ser	Ser	Ser 525	Ala	Gly	Ala	Arg	1667
	Trp 530	Arg	Glu	Ala	Leu	Pro 535	Gly	CCT Pro	His	Tyr	Tyr 540	Thr	Trp	Gly	Asp	His 545	1715
40	GGC Gly	GGA Gly	ATC Ile	ATC Ile	ACG Thr 550	GCC Ala	ATT Ile	GCC Ala	CAG Gln	GGC Gly 555	ATG Met	GAA Glu	ACC Thr	AAC Asn	GAG Glu 560	CTA Leu	1763
	Lys	Tyr	Ser	Thr 565	Asn	Glu	Gly	GAG Glu	Thr 570	Trp	Lys	Thr	Phe	Ile 575	Phe	Ser	1811
45	Glu	Lys	Pro 580	Val	Phe	Val	Tyr	GGC Gly 585	Leu	Leu	Thr	Glu	Pro 590	Gly	Glu	Lys	1859
45	Ser	Thr 595	Val	Phe	Thr	Ile	Phe 600		Ser	Asn	Lys	Glu 605	Asn	Val	His	Ser	1907
	Trp 610	Leu	Ile	Leu	Gln	Val 615	Asn	GCC Ala	Thr	Asp	Ala 620	Leu	Gly	Val	Pro	Cys 625	1955
50	Thr	Glu	Asn	Asp	Tyr 630	Lys	Leu	TGG Trp	Ser	Pro 635	Ser	Asp	Glu	Arg	Gly 640	Asn	2003
	GAG	TGT	TTG	CTG	GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	CGG	ACC	CCC	CAT	2051

				043	•				650					655		His	
	GCC	ACA	TGC	TTC	AAT	' GGA	GAG	GAC	TTTT	GAC	AGG	CCG	GTC	CTC	- Cmic	TCC	
_	Ala	Thr	Cys	Phe	Asn	Glv	Glu	Agn	Dhe	Agn	Ara	Dro	77-1	77-3	GIG	Ser	2099
5			660					665			9	FIO			val	Ser	
	AAC	TGC	TCC	TGC	ACC	CCC	GNG	GNO	יייאייי	C3.C	· · · · · · · · · · · · · · · · · · ·		670			AAG	
	Asn	Cva	Ser	Cvc	The	7~~	Clu	BAC	IMI	GAG	161	GAC	TTC	GGT	TTC	AAG	2147
	11011	675	JUL	Cys	1111	Arg	GIU	Asp	Tyr	GIU	cys	Asp	Phe	Gly	Phe	Lys	
	N TO			C1 m	mma		680					685					
	AIG	AGI	GAA	GAT	TTG	TCA	TTA	GAG	GTT	TGT	GTT	CCA	GAT	CCG	GAA	TTT	2195
10	Mec	ser	GIU	Asp	Leu	Ser	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Glu	Phe	
10	690					695					700		_			705	
	TCT	GGA	AAG	TCA	TAC	TCC	CCT	CCT	GTG	CCT	TGC	CCT	GTG	GGT	TCT	A CITI	2243
	Ser	Gly	Lys	Ser	Tyr	Ser	Pro	Pro	Val	Pro	Cvs	Pro	Val	Glv	Ser	Thr	-213
					710					715					720		
	TAC	AGG	AGA	ACG	AGA	GGC	TAC	CGG	AAG	ATT	TCT	ccc	GAC	аст	TOT	NCC.	2201
	Tyr	Arq	Arg	Thr	Ara	Glv	Tyr	Ara	Lvg	Tle	Ser	Gly	VCD	The	C	AGC	2291
15	-	_	_	725		2	- 2 -	3	730		501	Gry	App		Cys	ser	
	GGA	GGA	GAT		GAA	GCG	CGA	CTC	(A)	CCA	CNC	ama	C.T.C	735			
	Glv	Glv	Agn	Val	Glu	712	y~~	Tou	Cl	Clas	GAG	CIG	GTC	CCC	TGT	CCC	2339
	0+7	OL,	740	Val	Giu	nia	Arg	Leu	GIU	GIY	GIU	Leu		Pro	Cys	Pro	
	CTG	CCA		CAC	330	030	mm.a	745					750				
	tou	Ala	Clu	CAG	AAC	GAG	TTC	ATT	CTG	TAT	GCT	GTG	AGG	AAA	TCC	ATC	2387
	red	Ald	GIU	GIU	Asn	GIU	Phe	He	Leu	Tyr	Ala	Val	Arg	Lys	Ser	Ile	
20	m> a	755		~			760					765					
	TAC	CGC	TAT	GAC	CTG	GCC	TCG	GGA	GCC	ACC	GAG	CAG	TTG	CCT	CTC	ACC	2435
	Tyr	Arg	Tyr	Asp	Leu	Ala	Ser	Gly	Ala	Thr	Glu	Gln	Leu	Pro	Leu	Thr	
	770					775					780					705	
	GGG	CTA	CGG	GCA	GCA	GTG	GCC	CTG	GAC	TTT	GAC	TAT	GAG	CAC	AAC	TOT	2483
	Gly	Leu	Arg	Ala	Ala	Val	Ala	Leu	Asp	Phe	Asp	Tvr	Glu	His	Asn	Cve	2403
					790					795					800	_	
25	TTG	TAT	TGG	TCC	GAC	CTG	GCC	TTG	GAC	GTC	ATC	CAG	CGC	CTC	TOT	TTC	2531
	Leu	Tyr	Trp	Ser	Asp	Leu	Ala	Leu	Asp	Val	Tle	Gln	Ara	Lau	Cva	Tan	2531
		-	-	805					810	· u _	110	GIII	ALG	815	Cys	пеа	
	AAT	GGA	AGC		GGG	CAA	GAG	CTC	ATC	አጥር	ייית א	TOT	cca	012	~~	3.03	
	Asn	Glv	Ser	Thr	Glv	Gln	Glu	V-1	TIC	TI-	AAI	701	23	CIG	GAG	ACA	2579
		<b>-</b> - <i>y</i>	820		Gry	GIII	GIU	vai	TIE	ire	ASI	Ser		Leu	GIu	Thr	
30	CTA	CAA		<b>T</b> TO	CCT	TOTAL		825	~~~		~~~		830				
	Val	Clu	λla	Tou	772	Dho	GAA	CCC	CTC	AGC	CAG	CTG	CTT	TAC	TGG	GTA	2627
	Val	835	AIA	neu	Ата	Pne	Glu	Pro	Leu	Ser	GIn		Leu	Tyr	Trp	Val	
	CAT		000	mma			840					845					
	GWI	Bl-	93	TIC	AAA	AAG	ATT	GAG	GTA	GCT	AAT	CCA	GAT	GGC	GAC	TTC	2675
	ASP	Ala	GIA	Pne	Lys	rys	Ile	Glu	Val	Ala	Asn	Pro	Asp	Gly	Asp	Phe	
	850					855					860				-	965	
<i>35</i>	CGA	CTC	ACA	ATC	GTC	AAT	TCC	TCT	GTG	CTT	GAT	CGT	CCC	AGG	GCT	CTG	2723
	Arg	Leu	Thr	Ile	Val	Asn	Ser	Ser	Val	Leu	Asp	Arq	Pro	Ara	Ala	Leu	
					870					875					880		
	GTC	CTC	GTG	CCC	CAA	GAG	GGG	GTG	ATG	TTC	TGG	ACA	GAC	TGG	CCA	GAC	2771
	Val	Leu	Val	Pro	Gln	Glu	Gly	Val	Met	Phe	Tro	Thr	Asp	Trn	Glv	Aen	2//1
				885					890					895			
40	CTG	AAG	CCT	GGG	ATT	TAT	CGG	AGC	AAT	ATG	GAT	CCT	тСт	CCT	CCC	ጥለጥ	2010
40	Leu	Lys	Pro	Glv	Ile	Tvr	Arg	Ser	Asn	Met	Acn	Glv	Sor	אום	אום	Tres	2819
		-	900	- 4		- 1 -		905		1100	ASP	GIY		Ата	AId	TAT	
	CAC	CTG	GTG	ጥርጥ	GAG	САТ	GTG	703	TYC	ccc	2 2 10	<b>a</b> aa	910	mam	ama	~~~	
	His	Leu	Val	Ser	Glu	Agn	Val	Tiro	T	D~-	WWI	03	AIC	101	GIG	GAC	2867
		915			014	ASD	920	БÅВ	пр	PIO	Asn		TIE	ser	vaı	Asp	
	GAC		TCC	አ ጥጥ	The C	TOO	320	~~ m	~~~			925					
45	Acr	Cla	Too.	All Tla	TAC	166	ACG	GAT	GCC	TAC	CTG	GAG	TGC	ATA	GAG	CGG	2915
	930	GIII	IIP	тте	Tyr	Trp	Thr	Asp	Ala	Tyr	Leu	Glu	Cys	Ile	Glu	Arg	
		3.00	~~~			935					940					945	
	AIC.	ACG	TTC	AGT	GGC	CAG	CAG	CGC	TCT	GTC	ATT	CTG	GAC	AAC	CTC	CCG	2963
	TTE	Thr	Phe	Ser	Gly	Gln	Gln	Arg	Ser	Val	Ile	Leu	Asp	Asn	Leu	Pro	
					950					955					960		
	CAC	CCC	TAT	GCC	ATT	GCT	GTC	TTT	AAG	AAT	GAA	ATC	TAC	TGG	GAT	GAC	3011
50	His	Pro	Tyr .	Ala	Ile	Ala	Val	Phe	Lys	Asn	Glu	Ile	Tyr	Tro	Asp	Asp	
				965					970					975			
	TGG	TCA	CAG	CTC	AGC	ATA	TTC	CGA	GCT	TCC	AAA	TAC	AGT	GGG	ጥርር	CAG	3059
	Trp	Ser	Gln	Leu	Ser	Ile	Phe	Ara	Ala	Ser	Lve	Tvr	Ser	61 v	500	Clr	2022
	-				-	_		3			_, _	- 7 ~		~-y	JC1		

		980		985			940			
	Met Glu	Ile Leu	GCA AAC Ala Asn	Gln Leu		ly Leu	Met Asp			3107
5		AAG GGG	AAG AAC Lys Asn			AT GCC				3155
	1010		101 TGC CTG	5		1020	_		1025	3203
			Cys Leu 1030		Ala A			Ser C		
10		Glu Asp	GTG TCC Val Ser		· Val L			Asp I		3251
		104	:5 'CAG GGC		1050	30 330	***** ****	1055	ማር እእአ	3299
	Cys Asp	Cys Pro	Gln Gly	Tyr Gln	Leu L	ys Asn	Asn Thr	Cys V	al Lys	3299
15	GAA GAG	AAC ACC	TGT CTT			AT CGC	TGC AGC	AAC C	GG AAC	3347
	107	'5	Cys Leu	1080			1085			
	TGT ATC	AAC AGC	ATT TGG	TGG TGT	GAC T	TT GAC	AAC GAC	TGT	GA GAC	3395
	Cys Ile	Asn Ser	: Ile Trp 109		s Asp P	ne Asp 1100		Cys (	1105	
20	ATG AGO	GAT GAG	AGA AAC	TGC CCI	ACC A	CC ATC	TGT GAC	CTG (	SAC ACC	3443
		_	Arg Asr		1	.115			1120	2402
	CAG TTT	CGT TGC	CAG GAG	TCT GGG	ACT T	GT ATC	CCA CTG	TCC :	rar aaa	3491
		112			1130			1135		
25	TGT GAC	CTT GAC	GAT GAC	TGT GG	A GAC A	AC AGT	GAT GAA	AGT (	CAT TGT	3539
25	Cys Asp	Leu Glu 1140	a Asp Asp	Cys Gly 114		Asn Ser	Asp Glu 115		His Cys	
	GAA ATG	CAC CAC	TGC CGG	AGT GAG	GAG T	CAC AAC	TGC AGT	TCC (	GGC ATG	3587
	Glu Met		Cys Arg	Ser Asp 1160	Glu T		Cys Ser 1165	Ser (	Gly Met	
	TGC ATO	CGC TCC	TCC TGC		GAC G			TGC Z	AGG GAC	3635
30			Ser Tr	Val Cys		ly Asp	Asn Asp		Arg Asp	
	1170		117 A GCC AAC		- ccc »	1180	) מאר ארר	TOTO (	1185	3683
	TGG TCT	GAT GAZ	A GCC AAC 1 Ala Asr	CVS Th	r Ala I	lle Tvr	His Thr	Cvs	Glu Ala	2002
	Ξ,	_	1190		1	L <b>1</b> 95			1200	
•	TCC AAC	TTC CAC	TGC CGA	AAC GGO	G CAC I	GC ATC	CCC CAG	CGG '	TGG GCG	3731
35		120	n Cys Arg		1210			1215		
	TGT GAG	GGG GA	r acg gad	TGC CA	GAT C	GT TCC	GAT GAG	GAT	CCA GTC	3779
	Cys Ası	Gly Ası 1220	Thr Asp	Cys Gli 12:		Gly Ser	Asp Glu 123		Pro Val	
	AAC TG	GAG AA	AAG TG	C AAT GG	A TTC	CGC TGC	CCA AAC	GGC .	ACT TGC	3827
40	Asn Cys		s Lys Cys	s Asn Gly 1240	y Phe A	Arg Cys	Pro Asn 1245	Gly	Thr Cys	
	ATC CC	A TCC AG	C AAA CA	r TGT GA'	r GGT (	CTG CGT	GAT TGC	TCT	GAT GGC	3875
	Ile Pro	Ser Se	r Lys His	s Cys As	p Gly I	Leu Arg	Asp Cys	Ser .	Asp Gly	
	1250		12!	55	a ama i	1260		a mc	1265	3923
	TCC GA	r GAA CAG	G CAC TGG n His Cy:	GAG CC	o ten (	Ove Thr	His Phe	Met	Asp Phe	3723
45			1270			1275			1280	
	GTG TG	T AAG AA	C CGC CA	G CAG TG	C CTG	TTC CAC	TCC ATC	GTC	TGT GAC	3971
	Val Cy	s Lys As: 12	n Arg Gli	n Gln Cy	s Leu 1 1290	Phe His	Ser Met	Val 1295	cys Asp	
	GGA AT	C ATC CA	G TGC CG	C GAC GG	G TCC	GAT GAG	GAT GCG	GCG	TTT GCA	4019
50	Gly Il	e Ile Gl	n Cys Ar	g Asp Gl	y Ser A	Asp Glu	Asp Ala	Ala	Phe Ala	
50	GGA TO	1300 C TCC CA	A GAT CC	13 T GAG TT	US C C <u>A</u> C :	AAG GTA	131 TGT GAT		TTC GGT	4067
	Gly Cv	s Ser Gl	n Asp Pr	o Glu Ph	e His	Lys Val	Cys Asp	Glu	Phe Gly	
	13		-	1320			1325			

	TTC Phe	CAG Gln	TGT Cys	CAG Gln	AAT Asn	GGA Gly	GTG Val	TGC Cys	ATC Ile	AGT Ser	TTG Lei	ATT Ile	TCC TCC	AAG Lys	TGC Cvs	GAC Asp	4115
	1330		- 2 -			1335		- 4			1340			-3-	- 7 -	1345	
	GGG	ATG	GAT	GAT	TGC	GGC	GAT	TAT	TCT	GAT	GAA	GCC	AAC	TGC	GAA	AAC	4163
5	Gly	Met	Asp	Asp	Cys	Gly	Asp	Tyr	Ser	Asp	Glu	Ala	Asn	Cys	Glu	Asn	
	_				1350	)				1355	5				1360	)	
						AAC											4211
	Pro	Thr	Glu	Ala	Pro	Asn	Cys	Ser			Phe	Gln	Phe			Glu	
				1365	<b>i</b>				1370					1375			
_	AAT	GGC	CAC	TGC	ATC	CCC	AAC	AGA	TGG	AAA	TGT	GAC	AGG	GAG	AAC	GAC	4259
10	Asn	Gly			He	Pro	Asn			гÀа	Cys	Asp			Asn	Asp	
			1380		mam	GAT	CAC	1385		TOT	CCA	CAT	1390		አጥጥ	CTT	4207
	TGT	(3)	CAC	Tee	Co~	Asp	Clu	Luc	GWI	Cve	Glv	Aen	Ser	Uie	Tla	Len	4307
	Cys	1395		пр	Set	ASP	1400		АЗР	Cys	GLY	1405		1112	116	neu	
	CCC	TOPE	TCG	ΔСТ	CCT	GGG			ACG	TGT	CTG			TAC	TAC	CGC	4355
	Pro	Phe	Ser	Thr	Pro	Gly	Pro	Ser	Thr	Cvs	Leu	Pro	Asn	Tyr	Tvr	Arg	
15	1410					1415				-1-	1420			-		1425	
	TGC	AGC	AGT	GGG	ACC	TGC		ATG	GAC	ACC	TGG	GTG	TGC	GAC	GGG	TAC	4403
	Cys	Ser	Ser	Gly	Thr	Cys	Val	Met	Asp	Thr	Trp	Val	Cys	Asp	Gly	Tyr	
	-			_	1430	) _				1435	5				1440	)	
	CGA	GAT	TGT	GCA	GAT	GGC	TCT	GAC	GAG	GAA	GCC	TGC	CCC	TTG	CTT	GCA	4451
20	Arg	Asp	Cys	Ala	Asp	Gly	Ser	Asp	Glu	Glu	Ala	Cys	Pro			Ala	
				1445					1450					1455			
	AAC	GTC	ACT	GCT	GCC	TCC	ACT	CCC	ACC	CAA	CTT	GGG	CGA	TGT	GAC	CGA	4499
	Asn	Val			Ala	Ser	Thr			GIn	Leu	GIY			Asp	Arg	
			1460	)		~~~	<b>~~~</b>	1465		3.00	mam.	3.00	1470		maa	***	4547
	TTT	GAG	TTC	GAA	TGC	CAC	CAA	Doo	AAG	ACG	Cur	AIT	DTO	AAC	TGG	AAG	4547
25	Pne			GIU	Cys	His			гÃз	Int	Cys	1485		ASII	пр	пуs	
	000	1475	C 3 C	ccc	CAC	CAA	1480		CAG	CAT	ccc			GAG	GCC	דעע	4595
	750	Cra	AGN	GGC	LAC	Gln	Aen	Cve	Gln	Agn	Glv	Ara	Asp	Glu	Ala	Asn	4333
	1490		Asp	GLY	nis	1495		Cys	<b>J</b>	p	1500					1505	
	TGC	כככ	ACA	CAC	AGC	ACC		ACT	TGC	ATG			GAG	TTC	CAG		4643
	Cvs	Pro	Thr	His	Ser	Thr	Leu	Thr	Cys	Met	Ser	Arg	Glu	Phe	Gln	Cys	
30	_				1510	)				1519	5				1520	)	
	GAG	GAC	GGG	GAG	GCC	TGC	ATT	GTG	CTC	TCG	GAG	CGC	TGC	GAC	GGC	TTC	4691
	Glu	Asp	Gly	Glu	Ala	Cys	Ile	Val	Leu	Ser	Glu	Arg	Cys	Asp	Gly	Phe	
		_	_	1525	5				1530					1535			
	CTG	GAC	TGC	TCG	GAC	GAG	AGC	GAT	GAA	AAG	GCC	TGC	AGT	GAT	GAG	TTG	4739
	Leu	Asp	Cys	Ser	Asp	Glu	Ser			Lys	Ala	Cys			Glu	Leu	
35			1540					1545					1550				
	ACT	GTG	TAC	AAA	GTA	CAG	AAT	CTT	CAG	TGG	ACA	GCT	GAC	TTC	TCT	GGG	4787
	Thr			Lys	Val	Gln			GIn	Trp	Thr			Pne	ser	GIY	
	~~~	1555	· ~			maa	1560		000		***	1569		TOT	CCN	TOT	4835
	GAT	GIG	ACT	TTG	ACC	TGG	ATG	AGG	CCC	AAA	AAA	Mot	BTO	Ser	Ala	Ser	4033
	1570		Inr	Leu	Int	Trp 157		MIG	PIO	Lys	158	n Met	FIC	Ser	AIG	1585	
40			тат	አአጥ	GTC			AGG	GTG	بلبلت			AGC	ATA	TGG	AAG	4883
	Cve	V=1	TAL	VAL	Val	Tyr	TVY	Ara	Val	Val	Glv	Glu	Ser	Ile	Tro	Lvs	
	Cys	Val	- 7 -	AJII	159		- 7 -	9		159	5				160	0	
	ACT	CTG	GAG	ACC	CAC	AGC	AAT	AAG	ACA			GTA	TTA	AAA			4931
	Thr	Leu	Glu	Thr	His	Ser	Asn	Lvs	Thr	Asn	Thr	Val	Leu	Lys	Val	Leu	
				160				2	161					161			
45	AAA	CCA	GAT	ACC	ACG	TAT	CAG	GTT	AAA	GTA	CAG	GTT	CAG	TGT	CTC	AGC	4979
	Lys	Pro	Asp	Thr	Thr	Tyr	Gln	Val	Lys	Val	Gln	Val	Gln	Сув	Leu	Ser	
			1620	0				162	5				163	0			
	AAG	GCA	CAC	AAC	ACC	AAT	GAC	TTT	GTG	ACC	CTG	AGG	ACC	CCA	GAG	GGA	5027
	Lys	Ala	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly	
		163	5				164	0				164	5				
50	TTG	CCA	GAT	GCC	CCT	CGA	AAT	CTC	CAG	CTG	TCA	CTC	CCC	AGG	GAA	GCA	5075
			Asp	Ala	Pro	Arg		Leu	Gln	Leu	Ser	Leu	Pro	Arg	GIU	ALA	
	165					165					166		a. a	3.77	(TAID	1665 GGC	5123
	~ ~ ~ ~	CCT		יליידי ע	מימים	CCC	-c	באבאיני	12/7	יירייןיי	יירירי	ATT	· Al'	44 ((LAT	Latal.	2143

					TO \	U				167	Pro 5				160	^	
5	CTC Leu	ATC Ile	CGT Arg	GAG Glu	TAC Tyr	ATT Ile	GTA Val	GAA Glu	TAC	AGC	AGG Arg	AGT Ser	GGT Gly	TCC Ser	AAC	N-TO-C	5171
J	TGG	GCC	TCC	CAG	AGG	GCT	GCT	AGT	169	O TTT	ACA	GAA	ATC	169	5 	ጥ ጥ እ	5219
			170	U				170	5		Thr		171	0			F267
10	Leu	171	Asn 5	Thr	Leu	Tyr	Thr 172	Val 0	Arg	Val	Ala	Ala 172	Val 5	Thr	Ser	Arg	5267
	GGA Gly 1730	Ile	GGA Gly	AAC Asn	TGG	AGC Ser 173	Asp	TCT Ser	AAA Lys	TCC Ser	ATT Ile 1740	Thr	ACC Thr	ATA Ile	AAA Lys	Gly	5315
	AAA	GTG	ATC	CCA	CCA	CCA	GAT	ATC	CAC	ATT	GAC	AGC	TAT	GGT	GAA	1745 AAT	5363
15	Lys	Val	Ile	Pro	Pro 1750	Pro	Asp	Ile	His	Ile 1759	Asp	Ser	Tyr	Gly	Glu	Asn O	7005
	Tyr	Leu	Ser	Phe 176!	Thr 5	Leu	Thr	Met	Glu 1770	Ser	GAT Asp	Ile	Lys	Val	Asn 5	Gly	5411
	TAT Tyr	GTG Val	GTG Val 1780	Asn	CTT Leu	TTC Phe	TGG Trp	GCA Ala 1789	Phe	GAC Asp	ACC Thr	CAC His	AAG Lys 179	CAA Gln	GAG	AGG Arg	5459
20	AGA	ACT	TTG	AAC	TTC	CGA	GGA	AGC	ATA	TTG	TCA	CAC	AAA	GTT	GGC	AAT	5 507
	Arg	1795	5				1800	0				1809	5		_		
	CTG Leu 1810	Thr	Ala	Hıs	Thr	Ser 1819	Tyr	Glu	Ile	Ser	Ala 1820	Trp	Ala	Lys	Thr	Asp 1825	5555
25	TTG Leu	GGG Gly	GAT Asp	AGC Ser	Pro	Leu	GCA Ala	TTT Phe	GAG Glu	His	GTT Val	ATG	ACC Thr	AGA Arg	Gly	GTT Val	5603
	CGC Arg	CCA Pro	CCT Pro	Ala	Pro	AGC	CTC Leu	AAG Lys	GCC Ala	1835 AAA Lys	GCC	ATC Ile	AAC Asn	CAG Gln	1840 ACT Thr	GCA	5651
30	GTG	GAA	TGT	ACC	TGG	ACC	GGC	CCC	1850 CGG	AAT	GTG	GTT	TAT	1859 GGT	ATT	TTC	5699
	Val		1860)				1865	5				1870)			
	TAT Tyr	Ala	ACG Thr	TCC	TTT	CTT	GAC	CTC	TAT	CGC	AAC	CCG	AAG	AGC	TTG	ACT	5747
		1875	•				1880)				1885	;				
35	ACT	TCA	CTC	CAC	AAC	AAG	ACG	GTC	ATT	GTC	AGT	AAG	GAT	GAG	CAG	TAT	5795
	Thr 1890					1895	;				1900)				1905	
	TTG Leu	TTT Phe	CTG Leu	GTC Val	CGT Arg 1910	Val	GTG Val	GTA Val	CCC Pro	TAC Tyr 1915	Gln	GGG Gly	CCA Pro	TCC Ser	TCT Ser	Asp	5843
40	TAC Tyr	GTT Val	GTA Val	Val	AAG Lys	ATG	ATC Ile	CCG Pro	Asp	AGC Ser	AGG	CTT Leu	CCA Pro	CCC Pro	CGT	CAC	5891
	CTG Leu	CAT His	Val	Val	CAT	ACG Thr	GGC Gly	AAA Lys	1930 ACC Thr	TCC	GTG Val	GTC Val	ATC Ile	1935 AAG Lys	TGG	GAA Glu	5939
4 5	TCA Ser	Pro	Tyr	GAC	TCT Ser	CCT Pro	GAC Asp	1945 CAG Gln	GAC	TTG Leu	TTG Leu	TAT Tyr	1950 GCA Ala	ATT	GCA Ala	GTC Val	5987
	AAA	1955					1960)				1965	;				6025
	Lys 1970	Asp	Leu	Ile	Arg	Lys 1975	Thr	Asp	Arg	Ser	Tyr 1980	Lys	Val	Lys	Ser	Arg 1985	6035
50	AAC Asn	AGC Ser	ACT Thr	GTG Val	Glu	Tyr	ACC Thr	CTT Leu	AAC Asn	Lys	TTG Leu	GAG	CCT Pro	GGC Gly	Gly	AAA Lys	6083
	TAC Tyr	CAC His	ATC Ile	ATT Ile	1990 GTC Val	CAA	CTG Leu	GGG Gly	AAC Asn	1995 ATG Met	AGC	AAA Lys	GAT Asp	TCC Ser	2000 AGC Ser	ATA	6131

				2005	5				2010					3012			
	AAA	ATT	ACC	ACA	GTT	TCA	TTA	TCA	GCA	CCJ,	GAT	GCC	TTA	AAA	ATC	Α'nΑ	6179
	Lys	Ile	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	Ile	Ile	
5	-		2020)				2025	5				2030)			
	ACA	GAA	AAT	GAT	CAT	GTT	CTT	CTG	TTT	TGG	AAA	AGC	CTG	GCT	TTA	AAG	6227
	Thr	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys	
		2035	5	_			2040)				2045	5				
	GAA	AAG	CAT	TTT	AAT	GAA	AGC	AGG	GGC	TAT	GAG	ATA	CAC	ATG	TTT	GAT	6275
	Glu	Lys	His	Phe	Asn	Glu	Ser	Arg	Gly	Tyr	Glu	Ile	His	Met	Phe	Asp	
10	2050)				2055	5				2060)				2065	
	AGT	GCC	ATG	AAT	ATC	ACA	GCT	TAC	CTT	GGG	AAT	ACT	ACT	GAC	AAT	TTC	6323
	Ser	Ala	Met	Asn	Ile	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn	Phe	
					2070)				2075	5				2080)	
	TTT	AAA	ATT	TCC	AAC	CTG	AAG	ATG	GGT	CAT	AAT	TAC	ACG	TTC	ACC	GTC	6371
15	Phe	Lys	Ile	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Phe	Thr	Val	
15		_		2089	5				2090	0				209	5		
	CAA	GCA	AGA	TGC	CTT	TTT	GGC	AAC	CAG	ATC	TGT	GGG	GAG	CCT	GCC	ATC	6419
	Gln	Ala	Arg	Cys	Leu	Phe	Gly	Asn	Gln	Ile	Cys	Gly	Glu	Pro	Ala	Ile	
			2100)				210	5				211	0			
	CTG	CTG	TAC	GAT	GAG	CTG	GGG	TCT	GGT	GCA	GAT	GCA	TCT	GCA	ACG	CAG	6467
20	Leu	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Gly	Ala	Asp	Ala	Ser	Ala	Thr	Gln	
		2115	5				212	0				212	5				
	GCT	GCC	AGA	TCT	ACG	GAT	GTT	GCT	GCT	GTG	GTG	GTG	CCC	ATC	TTA	TTC	6515
	Ala	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	Ile	Leu	Phe	
	2131	1				213	5				214	0				2145	c= c=
	CTG	ATA	CTG	CTG	AGC	CTG	GGG	GTG	GGG	TTT	GCC	ATC	CTG	TAC	ACG	AAG	6563
25	Leu	Ile	Leu	Leu	Ser	Leu	Gly	Val	Gly	Phe	Ala	He	Leu	Tyr	Thr	Lys	
					215	0				215		~~~			216		cc11
	CAC	CGG	AGG	CTG	CAG	AGC	AGC	TTC	ACC	GCC	TTC	GCC	AAC	AGC	CAC	TAC	6611
	His	Arg	Arg			Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	HIS	Tyr	
				216	5				217	0		-	a. m	217		ccc	6659
30	AGC	TCC	AGG	CTG	GGG	TCC	GCA	ATC	TTC	TCC	TCT	فافاف	GAT	GAC	CIG	GGG	0000
	Ser	Ser			Gly	Ser	Ala	Ile	Phe	Ser	ser	GTA	Asp	Asp	rea	Gly	
			218	0				218					219		a 2 a	CITIC .	6707
	GAA	GAT	GAT	GAA	GAT	GCC	CCT	ATG	ATA	ACT	GGA	777	TCA	GAT	GAC	GTC	6/0/
	Glu			Glu	Asp	Ala	Pro	Met	IIe	Thr	GIĀ	Pne	ser	Asp	Asp	Val	
		219	5				220	0				220		m	mcm x	አክጥአ	6762
35	CCC	ATG	GTG	ATA	GCC	TGA	AAGA	GCT	TTCC	TCAC	TA G	AAAC	CAAA	T. GG	IGIA	AATA	0/02
		Met	Val	Ile	Ala												
	221	0												ma 3	CITIES C	~~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	6022
								GTTT	A TT	'T'TAA	AAGA	TGC	ACTT	TGA	GITG	CAATAT	6843
	GTT	TTTA	TTA	TATG	GGCC	AA A											0043

Claims

40

50

55

1. DNA having a nucleotide sequence as shown by Sequence ID No. 1.

- 2. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 2 and coded by the DNA of Claim 1.
- 3. DNA having a nucleotide sequence as shown by Sequence ID No. 5.
- 4. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 6 and coded by the DNA of Claim 3.
- 5. A recombinant vector comprising DNA as shown by Sequence ID No. 1 or 5 and a replicable vector.
- 6. Transformant cells which harbor the recombinant vector of Claim 5.

7	•	A method for the production of an LDL receptor analog protein comprising the steps of culturing the transformants of Claim 6 and collecting a polypeptide produced in the culture.
5		
10		
15		
20		
<i>25</i>		
30		
35		
40		
4 5		
50		
55		

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 773 290 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3: 08.07.1998 Bulletin 1998/28

(51) Int. Cl.⁶: **C12N 15/12**, C07K 14/705 // C12N15/70, C12N15/79

(43) Date of publication A2: 14.05.1997 Bulletin 1997/20

(21) Application number: 96116108.0

(22) Date of filing: 08.10.1996

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE

(30) Priority: 09.10.1995 JP 261440/95 24.04.1996 JP 102451/96

(71) Applicant: KOWA COMPANY, LTD.
Naka-ku Nagoya-shi Aichi-ken (むア っ)

(72) Inventors:

Saito, Yasushi
 Chiba-shi, Chiba (JP)

Iwasaki, Akio
 Tsuchiura-shi, Ibaraki (JP)

 Arai, Koichi Urawa-shi, Saitama (JP)

Yamazaki, Hiroyuki
 Higashimurayama-shi, Tokyo (JP)

(74) Representative:
Wächtershäuser, Günter, Prof. Dr.
Patentanwalt,
Tal 29
80331 München (DE)

(54) Novel LDL receptor analog protein and the gene coding therefor

(57) The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No.5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.



EUROPEAN SEARCH REPORT

Application Number EP 96 11 6108

	DOCUMENTS CONSIDER Citation of document with indice	Relevant	CLASSIFICATION OF THE	
ategory	of relevant passage	s	to claim	APPLICATION (Int.CI.6)
Ą	WO 94 10322 A (UNIVER * the whole document	SITY OF TEXAS)	1-7	C12N15/12 C07K14/705 //C12N15/70,
A	WO 95 13374 A (BAYLOF * the whole document	COLLEGE OF MEDICINE) 1-7	C12N15/79
A	KOZARSKY K. ET AL.: of low density lipopy deficiency in the War hyperlipemic rabbit wadenoviruses" JOURNAL OF BIOLOGICAL vol. 269, no. 18, 6 l pages 13695-13702, X * the whole document	rotein receptor tanabe heritable with recombinant L CHEMISTRY, May 1994, MD US, P002062181	1-7	
A	WILSON J.M. ET AL.: gene transfer in viv improvement of hyper density lipoprotein rabbits" JOURNAL OF BIOLOGICA vol. 267, no. 2, 15	o leads to transient cholesterolemia in lo receptor-deficient L CHEMISTRY,	ow	TECHNICAL FIELDS SEARCHED (Int.CI.6)
	pages 963-967, XP002 * the whole document	. * 		
l l	The present search report has t	peen drawn up for all claims		
	Place of search	Date of completion of the sean	1	Examiner
3	BERLIN	16 April 1998	<u> </u>	Panzica, G
ã A:	CATEGORY OF CITED DOCUMENTS particularly relevant if taken alone particularly relevant if combined with anot document of the same category technological background	E : sarlier pate after the fili ther D : document o L : document o	rinciple underlying int document, but p ing date cited in the applica cited for other reas the same patent for	tion ons